

Kar, S. P. et al. (2016) Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types. *Cancer Discovery*, 6(9), pp. 1052-1067. (doi:[10.1158/2159-8290.CD-15-1227](https://doi.org/10.1158/2159-8290.CD-15-1227))

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/121562/>

Deposited on: 28 April 2017



Published in final edited form as:

*Cancer Discov.* 2016 September ; 6(9): 1052–1067. doi:10.1158/2159-8290.CD-15-1227.

## Genome-wide Meta-analyses of Breast, Ovarian and Prostate Cancer Association Studies Identify Multiple New Susceptibility Loci Shared by At Least Two Cancer Types

*A full list of authors and affiliations appears at the end of the article.*

### Abstract

Breast, ovarian, and prostate cancers are hormone-related and may have a shared genetic basis but this has not been investigated systematically by genome-wide association (GWA) studies. Meta-analyses combining the largest GWA meta-analysis data sets for these cancers totaling 112,349 cases and 116,421 controls of European ancestry, all together and in pairs, identified at  $P < 10^{-8}$  seven new cross-cancer loci: three associated with susceptibility to all three cancers (rs17041869/2q13/*BCL2L1*; rs7937840/11q12/*INCENP*; rs1469713/19p13/*GATAD2A*), two breast and ovarian cancer risk loci (rs200182588/9q31/*SMC2*; rs8037137/15q26/*RCCD1*), and two breast and prostate cancer risk loci (rs5013329/1p34/*NSUN4*; rs9375701/6q23/*L3MBTL3*). Index variants in five additional regions previously associated with only one cancer also showed clear association with a second cancer type. Cell-type specific expression quantitative trait locus and enhancer-gene interaction annotations suggested target genes with potential cross-cancer roles at the new loci. Pathway analysis revealed significant enrichment of death receptor signaling genes near loci with  $P < 10^{-5}$  in the three-cancer meta-analysis.

### Keywords

breast cancer; ovarian cancer; prostate cancer; genome-wide association studies; pleiotropy

### Introduction

Breast, ovarian and prostate cancer are hormone-related cancers (1). Breast and ovarian cancer share several environmental and lifestyle risk factors that affect exogenous or endogenous estrogen exposure, while the androgens play a key role in the pathophysiology

Corresponding Author: Siddhartha P. Kar, Centre for Cancer Genetic Epidemiology, University of Cambridge, Strangeways Research Laboratory, Cambridge CB1 8RN, United Kingdom. Phone: 44-01223-761938; Fax: 44-01223-748628; sk718@medschl.cam.ac.uk.

<sup>24</sup>A list of members from the Australian Ovarian Cancer Study (AOCS) Study Group and the Australian Cancer Study (Ovarian Cancer) is provided in Acknowledgments

<sup>25</sup>A list of members from the Australian Prostate Cancer BioResource (APCB) BioResource is provided in Acknowledgments

<sup>109</sup>A list of members from the Kathleen Cunningham Foundation Consortium for research into Familial Breast cancer (kConFab) Investigators is provided in Acknowledgments

<sup>136</sup>A list of members from the Norwegian Breast Cancer Study (NBCS) Investigators is provided in Acknowledgments

<sup>168</sup>A list of members from the GENICA Network is provided in Acknowledgments

<sup>169</sup>A list of members from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium is provided in Acknowledgments

\*These authors contributed equally to this article.

The authors disclose no potential conflicts of interest.

of prostate cancer. Collectively, cancers at these three sites accounted for more than 420,000 new cases, or over 25% of all cancers diagnosed, in the United States in 2012 (2).

All three cancers are known to aggregate in the same families (3–5). The effects of a shared environment and of rare, highly penetrant alleles in established cancer predisposition genes explain only a part of the observed familial clustering (6). This suggests that there exist common, low-penetrance susceptibility variants with shared effects across these cancer types. Since 2007, genome-wide association studies (GWAS), replication studies, and a custom genotyping effort focused individually on breast, ovarian, and prostate cancers have identified multiple risk loci specific to each cancer (summarized in refs. (7–9)). The same studies have also identified three susceptibility loci where the most strongly associated single nucleotide polymorphism (SNP) or the index SNP exhibits pleiotropy and is common to two of the cancer types (10–15). Moreover, in several other regions, separate index SNPs for risk of at least two of the cancers are found in close proximity and the underlying signal may well be pleiotropic (16). However, genetic association studies for breast, ovarian, and prostate cancers to date have been designed to be cancer site-specific. Pleiotropy between them has not itself been leveraged as the basis for systematic genome-wide discovery of completely new cancer risk loci – loci that share an association with at least two, if not all three, of the cancers.

Given this background, we combined data from the largest and most recently published genome-wide association meta-analysis for susceptibility to breast cancer (7), ovarian cancer (8), and prostate cancer (9), in a single three-cancer meta-analysis of 228,770 individuals, and in pairwise combinations. We hypothesized that the substantial gain in power afforded by the cross-cancer meta-analyses would enable the identification, at genome-wide significance, of risk loci sharing an association with more than one of the three cancers that are novel for each of the cancers (17). Pleiotropic alleles at these loci may be modestly associated with each of the cancers and not previously detected at the standard threshold for genome-wide significance ( $P < 5 \times 10^{-8}$ ) in single-cancer studies due to sample size constraints. We also investigated whether the index SNP in regions so far known to contain associations with only one cancer out of breast, ovarian, or prostate cancer showed clear evidence for association with another cancer out of the three. Further, at the rare variant end of the genetic association spectrum, the identification of rare alleles that confer inherited susceptibility to multiple cancers in genes such as *BRCA1*, *BRCA2*, and *TP53* has yielded critical insights into their role in cancer etiology (18,19). Motivated by this observation, we annotated the new cross-cancer risk loci using cell-type specific expression quantitative trait locus (eQTL) and enhancer-gene interaction maps, and performed enrichment analysis for molecular pathways in the top regions spanning breast, ovarian, and prostate cancer to identify putative shared target genes and mechanisms potentially driving common genetic susceptibility across these hormone-related cancers.

## Results

### Study Populations

We used summary statistics for association with cancer risk from the largest and most recently published meta-analysis of GWA, replication and custom genotyping case-control

studies for each cancer (7–9). These meta-analyses included 62,533 women with breast cancer (including 12,412 women with estrogen receptor (ER)-negative tumors) and 60,976 controls, 15,437 women with invasive epithelial ovarian cancer (including 9,627 women with serous tumors) and 30,845 controls, and 34,379 men with prostate cancer and 33,164 controls. All individuals were of European ancestry and a total of 8,564 controls overlapped between the breast and ovarian cancer studies. The summary statistics were available for variants with minor allele frequency (MAF) > 0.5% that had either been genotyped or imputed independently in the breast, ovarian and prostate cancer studies using the 1000 Genomes Project (March 2012 release) European reference panel.

### New Associations with a Second Cancer at Known Single-cancer Risk Loci

We listed the published index SNP at each of the 92, 18 and 100 loci known to be associated with breast, ovarian and prostate cancer susceptibility, respectively, in European-ancestry populations (Supplementary Table S1). The list comprised 207 unique SNPs after accounting for the three SNPs that were each an index SNP for two cancers (rs10069690 at 5p15 and rs8170 at 19p13 for breast and ovarian cancer, and rs4245739 at 1q32 for breast and prostate cancer; refs. (10–15)). Separate index SNPs for two of the cancers were within 1 Mb of each other in 21 genomic regions, including two regions (at 6p22 and 8q24) that contained index SNPs for all three cancers (Supplementary Table S2).

Scanning for these 207 SNPs using the summary data identified novel associations with breast cancer susceptibility at the ovarian cancer index SNP rs635634 at 9q34/*ABO* ( $P_{\text{BrCa}} = 8.1 \times 10^{-7}$ ) and at the prostate cancer index SNPs rs6763931 at 3q23/*ZBTB38* ( $P_{\text{BrCa}} = 1.2 \times 10^{-6}$ ) and rs11214775 at 11q23/*HTR3B* ( $P_{\text{BrCa}} = 5.2 \times 10^{-5}$ ) with consistent direction of allelic effect between the previously reported and novel associations across cancer types (Table 1). Further, the risk (T) allele of the breast cancer index variant in *BRCA2*, rs11571833 (MAF = 0.8%), was associated with ovarian cancer risk ( $P_{\text{OvCa}} = 6.4 \times 10^{-8}$  for serous invasive ovarian cancer; odds ratios and additional details in Table 1) while the protective (T) allele of the breast cancer index SNP rs1830298 at 2q33/*ALS2CR12* was associated with prostate cancer risk ( $P_{\text{PrCa}} = 1.3 \times 10^{-6}$ ; Table 1). Thus, index SNPs at five loci so far known to be associated with only one cancer type demonstrated strong evidence for association with a second cancer type, out of breast, ovarian, and prostate cancer, at a significance level of  $6 \times 10^{-5}$  after Bonferroni correction for testing 207 SNPs in four ways (breast and ovarian cancer, breast and prostate cancer, ovarian and prostate cancer, and ER-negative breast and serous ovarian cancer). There was no known index SNP associated with the second cancer type within 1 Mb on either side of any of the new signals.

### Meta-analysis of Breast, Ovarian and Prostate Cancer Genome-wide Association Meta-analysis Data

Having examined the index SNPs at established risk loci for each cancer to uncover new cross-cancer association signals, we conducted a fixed-effects meta-analysis using the breast, ovarian and prostate cancer summary statistics for all variants that were nominally associated ( $P < 0.05$ ) with each of the three cancers. In effect, our study design enabled independent replication of findings reaching  $P < 0.05$  for association with susceptibility to one cancer type in data from the two other cancer types. We reasoned that this approach

could identify previously unrecognized cancer risk loci that were shared by breast, ovarian and prostate cancer and achieved genome-wide significance only after combining data from the three cancers.

The meta-analysis identified 267 alleles spanning 18 independent loci that were associated at  $P < 10^{-8}$  with breast, ovarian and prostate cancer susceptibility with the same direction of effect across all three cancers (Manhattan plot in Figure 1; Supplementary Table S3). The threshold for genome-wide significance was set at a more stringent  $P < 10^{-8}$  compared to the standard  $P < 5 \times 10^{-8}$  to correct for multiple comparisons arising from the fact that we searched for associations shared between cancer types in five possible ways (breast, ovarian and prostate cancer, breast and ovarian cancer, breast and prostate cancer, ovarian and prostate cancer, and ER-negative breast and serous ovarian cancer; the pairwise searches are described in the next section). Moreover, it is possible to obtain a genome-wide significant signal in a meta-analysis of three cancers in the setting of a particularly strong association between a variant and a subset of one or two of the cancers and no association with the remaining cancer(s). We addressed this possibility by applying the association analysis based on subsets (ASSET; ref. (20)) method to test whether the best association model for each newly identified index variant involved all three cancers, as would be expected for true cross-cancer signals. Model selection using ASSET demonstrated that all three cancers contributed to the signal at the most significantly associated variant at 13 of the 18 loci (Supplementary Table S3, ASSET column). None of these 13 index variants showed significant heterogeneity in the per-allele odds ratio between the three cancers further confirming consistent pan-cancer effects (Cochran's  $Q$ -test for heterogeneity,  $P_{\text{het}} > 0.05$ ).

To account for correlation between the breast and ovarian cancer studies due to the 8,564 controls shared between them, we repeated the meta-analysis for the 13 index variants using a statistical adjustment for studies with overlapping controls that required only summary statistics and exact sample counts contributing to the association at each variant from the corresponding data sets (21). Two of the variants fell just short while 11 remained at  $P < 10^{-8}$  after this adjustment (Supplementary Table S3,  $P_{\text{adjusted}}$  column). Eight of these 11 loci were less than 1 Mb from a known index SNP for at least one of the three cancers. Details of the eight susceptibility loci including linkage disequilibrium (LD) information with respect to known index SNPs are also presented in Supplementary Table S3.

The three remaining loci were over 1 Mb away from known index SNPs for any of the three cancers and indexed by the variants rs17041869 (in *BCL2L1* at 2q13;  $P_{\text{meta}} = 5.1 \times 10^{-9}$ ; Table 2A), rs7937840 (in *INCENP* at 11q12;  $P_{\text{meta}} = 5.0 \times 10^{-9}$ ), and rs1469713 (in *GATAD2A* at 19p13;  $P_{\text{meta}} = 3.4 \times 10^{-10}$ ). They represent entirely new association signals for all three cancers discovered at genome-wide significance ( $P < 10^{-8}$ ) by leveraging the shared genetic architecture of breast, ovarian and prostate cancer ( $P$ -values for each cancer type in Table 2A; regional association plots in Supplementary Fig. S1A–C). While the newly identified index variant rs1469713 itself was 960 kb from a known breast cancer index SNP rs4808801, 42 of the 89 variants in the new 19p13/*GATAD2A* region that were correlated with rs1469713 and reached  $P < 10^{-8}$  in the three-cancer meta-analysis were between 1 to 1.2 Mb away from rs4808801 (Supplementary Fig. S1C). Furthermore, rs1469713 and rs4808801 were not linked ( $r^2 = 0.001$  in the European populations from the 1000 Genomes

Project) and the association at rs1469713 remained on analysis of the breast cancer data conditioning on rs4808801, confirming independence of the new signal from the known one (Supplementary Table S4, which includes three-cancer meta-analysis results for rs1469713 undertaken using results from the conditional analysis). We also confirmed that the three new index SNPs were not correlated with known breast, ovarian or prostate cancer index SNPs up to 10 Mb away on either side ( $r^2 < 0.01$  in 1000 Genomes European populations). Figure 2A shows forest plots of odds ratios and 95% confidence intervals corresponding to association of the three novel index variants with each cancer separately and on meta-analysis. While rs17041869 had been genotyped, the two other index variants had been imputed with accuracy,  $r^2 = 0.89$  (Table 2A).

### Pairwise Meta-analyses using the Breast, Ovarian and Prostate Cancer Data

To identify new risk loci in common specifically to two of the three cancers, we combined data from the three cancer types in pairs using fixed-effects meta-analyses. We also conducted an additional meta-analysis for shared susceptibility to ER-negative breast and serous ovarian cancer as the two previously reported index SNPs known to be shared between breast cancer and ovarian cancer are specific to these two subtypes (rs10069690 and rs8170 (10–13)). Before examining results from each pairwise meta-analysis, we excluded all variants within 1 Mb of known index SNPs for either or both cancer types contributing to the meta-analysis to avoid detecting signals unduly driven by established associations in one or both cancer types contributing to the meta-analysis. We identified new shared associations with breast and ovarian cancer risk at rs200182588 (in *SMC2* at 9q31;  $P_{\text{meta}} = 8.9 \times 10^{-9}$  after adjusting for overlapping controls) and rs8037137 (near *RCCD1* at 15q26;  $P_{\text{meta}} = 9.1 \times 10^{-10}$  after adjustment), and with breast and prostate cancer risk at rs5013329 (in *NSUN4* at 1p34;  $P_{\text{meta}} = 1.8 \times 10^{-11}$ ) and rs9375701 (in *L3MBTL3* at 6q23;  $P_{\text{meta}} = 3.4 \times 10^{-10}$ ). Full results for the four new index SNPs are presented in Table 2B, forest plots in Figure 2B, and regional association plots in Supplementary Fig. 1D–G. These SNPs were not correlated with known index SNPs for the corresponding individual cancer types up to 10 Mb away on either side ( $r^2 < 0.01$  in 1000 Genomes European populations). ASSET confirmed contributions from both cancer types to each new signal, none of them displayed significant heterogeneity in the per-allele odds ratio ( $P_{\text{het}} > 0.05$ ), and the index variants had been imputed with accuracy,  $r^2 = 0.81$  (Table 2B). No new locus was identified at genome-wide significance ( $P < 10^{-8}$ ) in the ovarian and prostate cancer and in the subtype-specific ER-negative breast and serous ovarian cancer meta-analyses after excluding all variants within 1 Mb of known index SNPs for either or both cancer types contributing to the corresponding meta-analyses.

Further, there is some evidence that alleles that increase risk of one cancer confer protection from another cancer (notably at rs4245739 which is a known index SNP for both breast and prostate cancer). Therefore, we used the two-sided subset function implemented in ASSET to also look for alleles in the three-cancer meta-analysis that were associated with all three cancers (with a combined ASSET  $P < 10^{-8}$ ) but where the direction of allelic effect on one of the cancers was opposite to that observed for the other cancer types (details in Methods). To search for such alleles in each pairwise meta-analysis, we reversed the signs on the effect size estimates in one of the two data sets and repeated fixed-effects meta-analysis (22).



However, no novel loci were identified at  $P < 10^{-8}$  in the search for shared alleles with opposite effects on risk of different cancer types out of the three cancers using either approach.

Of the seven new loci identified by the three-cancer and pairwise meta-analyses, the index SNP at 2q13 is a genotyped SNP while the index SNPs at the remaining loci were well imputed with accuracy,  $r^2 = 0.81$  and had MAF = 12%. Imputation was conducted independently for data from each cancer type and imputation accuracy estimates were consistent across cancer types. In each of the three single-cancer genome-wide association meta-analyses that contributed to this three-cancer study, we have demonstrated high concordance between imputed and genotyped SNP results for common SNPs (MAF > 5%) identified at standard genome-wide significance ( $P < 5 \times 10^{-8}$ ) that have imputation accuracy > 0.80. Finally, for four of the six new loci where the index SNP was an imputed SNP, we were also able to identify a genotyped SNP in the same region that was also genome-wide significant ( $P < 10^{-8}$ ; Supplementary Table S5 provides results for the most significantly associated genotyped SNP at each of these six loci).

### Expression QTL Analyses Suggest Target Genes Shared Across Relevant Cell Types at New Loci

We carried out *cis*-expression quantitative trait locus (eQTL) analyses for the seven new index variants (listed in Table 2) and all genes up to 1 Mb on either side of each variant using breast (n = 183), ovarian (n = 85), and prostate (n = 87) normal tissue samples from the Genotype-Tissue Expression (GTEx) Project (23). The risk (T) allele of the breast and prostate cancer index variant rs9375701 was significantly associated with reduced expression of *L3MBTL3* in both breast ( $P = 3.5 \times 10^{-9}$ ) and prostate ( $P = 8.9 \times 10^{-7}$ ) tissue (box plots in Supplementary Fig. S2A). There were no significant associations ( $P < 0.05$ ) between this SNP and the expression level of any other gene in the same region in either cell type. A consistent cross-cell type association was also observed between the risk (T) allele of the breast and ovarian cancer index variant rs8037137 and decreased expression of *RCCD1* in both breast ( $P = 1.1 \times 10^{-15}$ ) and ovarian ( $P = 1.1 \times 10^{-5}$ ) tissue (box plots in Supplementary Fig. S2B). Some of the index variants also yielded eQTL associations that were nominally significant in only one of the three cell types (full results in Supplementary Table S6). Further, we looked up two of the seven index variants that were reported in a large database of eQTLs from peripheral blood samples (n = 5,311; ref. (24)) and found an association between the risk (G) allele of the three-cancer index SNP rs1469713 and increased expression of *GATAD2A* ( $P = 9.8 \times 10^{-198}$ ) and replicated the association between the T allele of rs9375701 and decreased expression of *L3MBTL3* ( $P = 4.7 \times 10^{-125}$ ).

### Cell-type Specific Enhancer Maps Suggest Target Genes Shared Across Relevant Cell Types at New Loci

Expression QTL analysis may not always be able to detect functionally important variant-gene relationships over background noise given small sample sizes of eQTL data sets, the dynamic nature of gene expression, and the likely modest biological effects of risk variants (25). Disease-associated genetic variation has been found enriched in cell-type specific

enhancer elements. Therefore, as an alternative strategy to identify potential cross-cancer susceptibility genes, we annotated all variants with  $P < 10^{-8}$  at the seven new loci (variant counts in Table 2) using maps of enhancers in breast, ovarian and prostate cell types (26–28). We intersected these maps with computationally predicted enhancer-gene interactions in the same cell types (26–28) as well as experimentally derived interactions that were only available for breast cells (HMEC and MCF7 profiled using Hi-C and ChIA-PET, respectively; refs. (29,30)).

Two intronic SNPs in *GATAD2A* out of the 89 variants with  $P < 10^{-8}$  at the 19p13 three-cancer risk locus (rs2916068 and rs2965183; Figure 3; Supplementary Table S7) were located in enhancers in normal and cancerous breast (HMEC and MCF7, respectively), normal ovarian (ovary, UCSD) and prostate cancer (LNCaP) cells and in each instance, this enhancer was predicted to interact with *GATAD2A*. A direct physical interaction between rs2916068 and the *GATAD2A* promoter was additionally confirmed in the MCF7 breast cancer cells assayed by ChIA-PET. The index SNP rs17041869 at the 2q13 three-cancer risk locus mapped to an enhancer in breast (MCF7) and prostate (LNCaP) cancer cells and in both cases, this enhancer was predicted to interact with *BCL2L11* (Figure 4; Supplementary Table S7). Notably, while evidence for *BCL2L11* as a target gene was not found in ovarian cells by enhancer-gene interaction annotation, the eQTL analysis did show a marginally significant association between the risk-conferring (G) allele of rs17041869 and elevated expression of *BCL2L11* in normal ovarian tissue samples ( $P_{\text{ovary\_eQTL}} = 0.048$ ), this being the only significant *cis*-association detected for rs17041869 in any of the three cell types (Supplementary Table S6).

At the 1p34 breast and prostate cancer region, several of the 218  $P < 10^{-8}$  variants overlapped enhancers that interacted with *NSUN4* in breast (MCF7; interaction confirmed by ChIA-PET) and prostate (LNCaP) cancer cells (Supplementary Table S7). SNP rs17361950 intersected enhancers interacting with *FAAH* in MCF7 (confirmed by ChIA-PET) and LNCaP while the indels chr1:46505589:I and chr1:46505785:I intersected enhancers targeting *PIK3R3* in LNCaP and in the breast cancer cell line HCC1954 (Supplementary Table S7). The risk (T) allele of the index SNP rs5013329 at the same locus was significantly associated with lower expression of *NSUN4* in breast ( $P_{\text{breast\_eQTL}} = 0.001$ ) and the long non-coding RNA (lncRNA) gene *RPS15AP10* in prostate ( $P_{\text{prostate\_eQTL}} = 0.02$ ) normal tissues (Supplementary Table S6). These findings collectively implicate *NSUN4* as the strongest shared functional candidate at 1p34. In addition to eQTL and enhancer mapping, we also annotated  $P < 10^{-8}$  variants in the seven regions using the HaploReg (31), lncRNASNP (32), and PolymiRTS (microRNA and miRNA target region annotation; ref. (33)) databases (Supplementary Table S8).

### Pathway Analysis Implicates Apoptosis as a Potential Mechanism for Susceptibility to All Three Cancers

Finally, we used pathway analysis to explore the genome-wide significant regions and the fraction of associations just failing to reach this threshold in the meta-analysis of data from the three cancers. We took all alleles (regardless of proximity to known index SNPs for any of the three cancers) that met three criteria: (i)  $P < 10^{-5}$  in the three-cancer meta-analysis,



(ii) same direction of effect across all three cancers, and (iii) no significant heterogeneity in the per-allele odds ratio between cancers ( $P_{\text{het}} > 0.05$ ). These 884 alleles were then subjected to LD-based ‘pruning’ to leave 69 independent alleles (details in Methods). Taking regions up to 1 Mb on either side of the 69 alleles and merging overlapping regions yielded 51 intervals harboring a shared association with breast, ovarian and prostate cancer at  $P < 10^{-5}$ . We used the Interval Enrichment tool (INRICH; ref. (34)) to permute 5,000 matched intervals and tested for enrichment of pathways from four databases (KEGG, Biocarta, Reactome, and Gene Ontology), correcting for multiple comparisons separately in each database. Only one pathway, from Biocarta, survived this correction: 8/32 genes from the induction of apoptosis through DR3 and DR4/5 Death Receptors signaling pathway (*CASP9*, *LMNA*, *CASP7*, *TNFSF10*, *TNFRSF10A*, *TNFRSF10B*, *RELA*, and *FADD*; ref. (35)) were located in 7/51 intervals (INRICH analysis  $P_{\text{empirical}} = 0.0004$ ,  $P_{\text{corrected}} = 0.01$ ; top SNP in each interval listed in Supplementary Table S9). *BCL2L1* – the likely target of the new 2q13 three-cancer risk locus – is not a member of this pathway but given that this gene is a known apoptosis facilitator (36,37), we also checked for and found interactions between *BCL2L1* and several members of the Biocarta Death Receptor signaling pathway (Supplementary Fig. S3; details in Methods). Moreover, other apoptosis-related pathways that did contain *BCL2L1* were among the top pathways in the Reactome and Gene Ontology databases ( $P_{\text{empirical}} = 0.04$ – $0.006$ ; Supplementary Table S10).

## Discussion

Here we report findings from the first cross-cancer type genome-wide association meta-analysis focused on three hormone-related cancers. Performing a series of fixed-effects meta-analyses to cover all possible combinations of these three cancers, and a subtype-specific analysis for ER-negative breast and serous ovarian cancers, we identified three loci demonstrating shared association with breast, ovarian, and prostate cancer risk, two with breast and ovarian cancer risk, and two with breast and prostate cancer risk. Each of these seven loci was over 1 Mb away from previously identified risk loci and had the same direction of allelic effect for the corresponding individual cancer types. They were followed up using cell-type specific eQTL and enhancer data to identify the gene(s) likely to be targeted by the risk variants that are in common across cell types. Although we prioritized discovery of cross-cancer risk loci that were novel for each of the cancers, we also found that the index SNP in five additional regions previously known to be associated with only one of the three cancers showed robust evidence for pleiotropic association with a second cancer type out of the three. Only one of these five showed opposite effects on the risk of two cancer types (rs1830298, a known breast cancer index SNP at 2q33, found to be associated with prostate cancer) possibly reflecting tissue-specific regulatory mechanisms and/or tissue-specific modulation by environmental factors at this locus.

Annotation of the new 19p13 three-cancer susceptibility locus revealed that two strongly associated variants ( $P < 10^{-8}$ ) intersected overlapping enhancer elements interacting with *GATAD2A* in breast, ovarian, and prostate cell types. *GATAD2A* is a subunit of the nucleosome remodeling and histone deacetylase (NuRD) complex, a chromatin-level regulator of transcription with a number of important and emerging roles in cancer biology (38). At the level of transcription, the NuRD complex is recruited by tissue-specific

oncogenic transcription factors to repress the expression of tumor suppressor genes while at the post-translational level, this complex has been shown to deacetylate p53 to inactivate p53-induced apoptosis. The index variant at the 2q13 three-cancer risk locus was located in enhancers targeting the apoptosis facilitator *BCL2L1* in breast and prostate cancers cells and was associated with expression of the same gene in normal ovarian cells (36,37). Interestingly, this variant rs17041869 is 53 kb away from rs6738028, a genome-wide significant index SNP for serum dehydroepiandrosterone sulphate (DHEAS) concentrations (39). Correlation between the two variants was  $r^2 = 0.08$  and  $D' = 1$ , with the cancer risk-conferring G allele (frequency = 0.10) of rs17041869 always segregating with the C allele (frequency = 0.59) of rs6738028 in 1000 Genomes European populations (although we did not find rs6738028 itself to be associated with cancer risk). Secreted largely by the adrenal glands, DHEAS is the most abundant circulating steroid in the human body and is converted into active androgens and estrogen in the relevant peripheral tissue (40,41). DHEAS levels have previously been linked to increased risk of breast cancer but the direction of its associations with cancers of the prostate and ovary are less clear (42–44). The DHEAS GWAS also showed that the C allele of rs6738028 was associated with higher DHEAS levels and significantly lower expression of *BCL2L1* in blood and adipose tissue, in keeping with the anti- and pro-apoptotic roles of DHEAS and *BCL2L1* (45,46), respectively. Taken together, these observations suggest that though independent variants may underlie the DHEAS and hormone-related cancer susceptibility signals at 2q13, the effects of both may be regulated through *BCL2L1*. While we were unable to highlight a particular target gene at the 11q12 three-cancer risk locus, the index variant and many linked SNPs lie in *INCENP* (Supplementary Fig. S1B) and the locus also includes *MTA2* (467 kb from the index variant), another member of the NuRD complex (38). *INCENP* codes for the inner centromere protein, a non-enzymatic subunit of the chromosomal passenger complex (CPC; ref. (47)), that serves as the scaffold for CPC assembly (48). The CPC is a master regulator of mitosis and the inner centromere protein is essential for the activation and cellular localization of the enzymatic subunit of the CPC, Aurora B kinase (49), which is a much-studied target with roles in multiple cancers (50). We have previously identified association between other correlated variants in *INCENP* and breast cancer susceptibility in a candidate gene study of CPC components though these associations did not reach genome-wide significance (51), further underscoring the utility of combining data across cancers to pick up far more robust signals.

Quantitative trait locus analysis identified a highly significant and directionally consistent cross-tissue association with *L3MBTL3* expression for the 6q23 breast and prostate cancer index SNP. *L3MBTL3* is a member of the malignant brain tumor (MBT) family of chromatin-modifying transcriptional repressors with histone code reading functions (52). Similarly, the eQTL data strongly suggested that *RCCD1* was a shared cancer susceptibility gene at the 15q26 breast and ovarian cancer risk locus. It is worth noting that the index variant at this locus, rs8037137, is correlated with rs2290203 ( $r^2 = 0.6$ ,  $D' = 1$  in 1000 Genomes European populations), which is a genome-wide significant index SNP for breast cancer predisposition in East Asians ( $P_{rs2290203} = 1.8 \times 10^{-6}$  in our breast-ovarian meta-analysis with same direction of effect as the East Asian breast cancer-specific signal; ref. (53)). SNPs in this region have not previously been associated with breast cancer risk in

Europeans or with ovarian cancer risk in any population. EQTL analysis in the East Asian study also identified the poorly characterized *RCCD1* as the likely target gene of the locus. A combination of enhancer and eQTL mapping implicated *NSUN4* as a potential breast and prostate cancer risk gene at 1p34. *NSUN4* encodes a methyltransferase with an important role in mitochondrial ribosome production (54). The index variant (rs200182588) at the 9q31 breast and ovarian cancer susceptibility locus lies in the 5'-untranslated region of *SMC2* and binds several transcription factors in diverse tissue types including c-Myc in MCF7 cells (Supplementary Table S8). The structural maintenance of chromosomes protein-2 encoded by *SMC2* is a core component of the condensin complex that is responsible for close packaging of chromatin before cell division (55). Moreover, *SMC2* is a direct transcriptional target of oncogenic WNT signaling and N-Myc (56,57), and is emerging as a critical player in the DNA damage response (58). The index risk SNP at one of the seven new cross-cancer susceptibility loci discussed here was a genotyped SNP while at four other loci we were able to identify a genotyped SNP in the same region that was also genome-wide significant at  $P < 10^{-8}$ . For the two remaining loci (9q31 and 11q12), the index SNPs were imputed SNPs (imputation accuracy  $> 0.8$ ) and should be followed-up with confirmatory genotyping in additional samples.

Pathway analysis indicated significant involvement of induction of apoptosis through DR3 and DR4/5 death receptor (DR) signaling in mediating global susceptibility to these three hormone-related cancers (35). In particular, our analysis revealed that 1 Mb windows around two SNPs on chromosomes 3 and 8 associated just short of genome-wide significance in the three-cancer meta-analysis ( $P_{rs3819772} = 7.6 \times 10^{-8}$  and  $P_{rs10113131} = 9.5 \times 10^{-7}$ ; Supplementary Table S9), harbored *TNFSF10* that codes for the TNF-related apoptosis-inducing ligand (TRAIL) and *TNFRSF10A* and *TNFRSF10B* that encode the two receptors of TRAIL, DR4 and DR5, respectively. DR5 expression in prostate cancer cells is androgen dependent and elevated levels of androgens have been shown to inhibit TRAIL-induced apoptosis in LNCaP (59). Likewise, most breast and ovarian cancer cell lines are resistant to TRAIL-induced apoptosis (60,61), likely due to estrogenic regulation of death receptor signals (62), endocytosis of cell surface DR4 and DR5 in breast cancer (63), and aberrant cleavage of the caspases in ovarian cancer (61). Given that recombinant TRAIL and its receptor agonist antibodies are already under development (64,65), the possible contribution of this druggable pathway to the risk of multiple hormone-related cancers might offer new avenues for early-stage cancer therapy.

In conclusion, we have demonstrated that pleiotropy or association of the same variant with multiple phenotypes, a genetic phenomenon recognized as early as Mendel's classic 1866 paper (66), can be tapped to combine genome-wide association data across cancer types and uncover several risk loci that are shared by – and represent novel findings for – breast, ovarian, and prostate cancer. Our preliminary *in silico* characterization of the new loci also suggests that the integration of orthogonal resources such as eQTL and enhancer annotations from different cell types enabled by cross-cancer site strategies may refine the post-GWAS identification of putative functional target genes at cancer risk loci (67). Finally, the increased power of pleiotropy-informed locus discovery, fine mapping, pathway analysis, and polygenic risk prediction over conventional single-cancer approaches has the potential to offer fresh insights into the common biology that may underpin susceptibility to these

three hormone-related cancers, with implications for cross-cancer genetic screening (68). This work thus illustrates the need for even larger pan-cancer genome-wide association meta-analyses that include data from a broad range of cancer types including the other hormone-related cancers.

## Methods

### Breast, Ovarian and Prostate Cancer Data Sets

The data sets contained SNP-level summary statistics from association analyses for cancer risk from a published meta-analysis of genome-wide association study (GWAS) discovery, replication, and custom genotyping case-control studies for each cancer. The relevant local institutional review board approved each of these studies, informed consent was obtained from participants, and the studies were conducted in accordance with the Declaration of Helsinki. Details of the study participants, genotyping, quality control, imputation, association analysis and meta-analysis for each data set have been previously described (7–9). All analyses in the current study were restricted to data from individuals of European ancestry. Genotypes in each data set had been imputed into the March 2012 release of the 1000 Genomes Project European ancestry reference panel (version 3 of the Phase 1 integrated variant set release; ref. (69)). We only considered results for variants imputed with imputation accuracy,  $r^2 > 0.3$ . Imputation accuracy estimates were calculated in samples from the Collaborative Oncological Gene-Environment Study (COGS; ref. (70)) since they comprised the largest subset of each data set. In addition to summary statistics for association with susceptibility to overall breast cancer, all invasive epithelial ovarian cancer, and overall prostate cancer, we also used summary results for association with estrogen receptor-negative breast cancer and serous epithelial ovarian cancer risks.

Compilation of the list of published breast, ovarian, and prostate cancer index SNPs is described under Supplementary Methods.

### Meta-analysis

Estimated magnitudes of association (beta coefficients) and standard errors for variants from each data set were combined assuming fixed effects using inverse-variance weighted meta-analysis implemented in METAL (71). Heterogeneity in the per-allele odds ratio between cancers was assessed using  $P$ -values from Cochran's  $Q$ -statistic also calculated by METAL. All linkage disequilibrium (LD) calculations ( $r^2$  and  $D'$ ) presented were performed using the LDlink suite and data from the 1000 Genomes Project European ancestry populations (69,72).

### Alleles with Opposite Effects and Contribution of Each Data Set to Association Signals

To identify alleles that confer risk of one cancer but are protective for another cancer in each two-cancer meta-analysis, we reversed the signs on the beta coefficients in one of the two data sets and repeated the corresponding meta-analysis as done previously by Zhernakova et al. (ref. (22)). To identify alleles sharing associations with breast, ovarian, and prostate cancer risk where the direction of allelic effect on any one of these cancers was opposite to that observed for the other two, we used the two-sided subset search function in the

association analysis based on subsets (ASSET) R package (version 1.0.0) (20). Specifically, we used the `h.traits` function with arguments set as follows: `side=2`, `meth.pval=c("DLM")`, and `search=2`. This function in ASSET searches for such alleles for subsets of data sets (in this case representing phenotypes or cancers) and calculates fixed-effects meta-analysis-style test statistics separately for each subset (that is, for two cancers in one direction and the third cancer in the opposite direction). The two test statistics are then combined using a Chi-squared statistic and corrected for the multiple subset searches conducted. We also used the model-selection function in ASSET to identify the subset of data sets or cancer types in each meta-analysis that contributed to the overall association signal at the newly identified index variants.

### Overlapping Controls Between the Breast and Ovarian Cancer Studies

An overlap of 8,548 controls existed between the breast and ovarian cancer data sets. To account for correlation between the data sets due to overlapping controls, we applied a general statistical decoupling framework that involves adjusting the standard errors of each variant from the dependent data sets using a correlation matrix generated from the sample overlap counts (21). The data sets can then be analyzed as independent data sets. The correlation matrix itself was calculated as previously described (73). Correlations between the overall breast and all invasive epithelial ovarian cancer data sets and between the ER-negative breast and serous ovarian cancer data sets were found to be ~8% and ~4%, respectively. We applied the decoupling framework using exact counts for samples contributing to the association at the index variant in each new region identified at  $P < 10^{-8}$  in any meta-analysis involving both the breast and ovarian cancer studies and repeated the corresponding meta-analysis using METAL to confirm the signal for the variant after adjustment of standard errors.

### Expression QTL Analyses

Expression QTL analysis results for each index variant at the seven loci listed in Table 2 and all genes within 1 Mb of it were looked up using the Genotype-Tissue Expression (GTEx) project Portal in normal breast ( $n = 183$ ), ovarian ( $n = 85$ ), and prostate ( $n = 87$ ) tissue samples (23). To improve the power to detect significant eQTLs, at the cost of losing tissue-specificity, we also performed the same searches (where data availability permitted) in the blood eQTL browser that is based on eQTL analysis in peripheral blood samples from 5,311 individuals (24).

### Enhancer-Gene Interactions

Maps of enhancer regions with predicted target genes were obtained from Hnisz et al. (26), Corradin et al. (28), and He et al. (27). Enhancers active in the breast cell types HMEC (normal), MCF-7 (cancer), and HCC1954 (cancer), normal ovarian cell types from UCSD, and the prostate cancer cell line LNCaP (as relevant to each locus) were intersected with all variants with  $P < 10^{-8}$  in the seven regions listed in Table 2 using Galaxy. ENCODE Chromatin Interaction Analysis by Paired-End Tag sequencing (ChIA-PET) data from MCF-7 cells (mediated by RNA polymerase 2 and ER $\alpha$ ) were downloaded using the UCSC Table browser (30). Galaxy was used to identify the ChIA-PET interactions between an implicated breast cell enhancer (containing a strongly associated variant) and a predicted

gene promoter (defined as regions 3 kb upstream and 1 kb downstream of the transcription start site).

### Functional Annotation

We annotated all variants with  $P < 10^{-8}$  at the seven loci listed in Table 2 using the HaploReg v3 pipeline (31). Variants were annotated with (a) their location within a gene or distance from the nearest gene, (b) their functional consequence as per dbSNP if they were intragenic (intronic; located in the 3' - or 5' -untranslated region; exonic: synonymous or nonsynonymous), (c) GERP and SiPhy conservation scores (74,75), (d) effect on regulatory (transcription factor binding) motifs calculated using position weight matrices obtained from TRANSFAC (76), JASPAR (77), and other sources(78), and (e) transcription factor binding data from ENCODE (78). We also annotated these SNPs based on whether they were located in long non-coding RNAs and microRNAs or microRNA seed regions and target sites (32,33). Regional association plots that integrated 1000 Genomes LD data with gene annotation tracks were generated using LocusZoom (79).

### LD-based Pruning and Pathway Analysis

All (884) alleles demonstrating the same direction of effect across cancers without significant heterogeneity in the per-allele odds ratio ( $P_{\text{het}} > 0.05$ ) and with association  $P$ -values  $< 10^{-5}$  in the meta-analysis of the three cancers were subjected to LD-based 'pruning' (80). Starting with the most significantly associated variant, all variants within 1 Mb of it with correlation,  $r^2 > 0.1$  (calculated using 1000 Genomes Project European population genotype data) were removed. This was followed by a second round of LD-pruning with the same  $r^2$  threshold but for a distance of 10 Mb to remove variants in long-range LD. This yielded 69 independent variants. Assuming that a variant could potentially regulate any gene up to 1 Mb on either side of it (81), we generated 69 2-Mb-wide intervals such that each was centered on one variant. Merging overlapping intervals left 51 intervals.

The Interval Enrichment (INRICH; ref. (34)) tool was used to permute 5,000 sets of intervals with each set reasonably well-matched to the original set of 51 intervals in terms of interval size, number of genes and variants per interval, and variant positions (sampled based on hg19 gene and 1000 Genomes variant location data). The permuted sets were used to calculate an empirical  $P$ -value for enrichment of genes from a particular pathway among the observed intervals. A second permutation step (1,000 permutations) was applied to correct for multiple comparisons at the pathway level. All pathways containing between 20 and 200 genes from four extensively-curated online pathway repositories: Biocarta, the Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and Gene Ontology were obtained from the Molecular Signatures Database (MSigDB v3.0; (82)). Four pathway databases were used because each has a distinct and largely complementary approach to capturing known biological pathways (83). However, considerable overlap was present in gene content of the common pathways across databases and therefore we applied INRICH separately to pathways from each database. The different types of biological interactions shown in Supplementary Figure 3 between *BCL2L1* and the genes in the Biocarta induction of apoptosis through DR3 and DR4/5 Death Receptors pathway were identified using the GeneMania server (84).



## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Authors

Siddhartha P. Kar<sup>1,\*</sup>, Jonathan Beesley<sup>2,\*</sup>, Ali Amin Al Olama<sup>1,\*</sup>, Kyriaki Michailidou<sup>1,\*</sup>, Jonathan Tyrer<sup>3,\*</sup>, ZSofia Kote-Jarai<sup>4</sup>, Kate Lawrenson<sup>5</sup>, Sara Lindstrom<sup>6</sup>, Susan J. Ramus<sup>5</sup>, Deborah J. Thompson<sup>1</sup>, ABCTB Investigators<sup>7</sup>, Adam S. Kibel<sup>8</sup>, Agnieszka Dansonka-Mieszkowska<sup>9</sup>, Agnieszka Michael<sup>10</sup>, Aida K. Dieffenbach<sup>11,12</sup>, Aleksandra Gentry-Maharaj<sup>13</sup>, Alice S. Whittemore<sup>14</sup>, Alicja Wolk<sup>15</sup>, Alvaro Monteiro<sup>16</sup>, Ana Peixoto<sup>17</sup>, Andrzej Kierzek<sup>10</sup>, Angela Cox<sup>18</sup>, Anja Rudolph<sup>19</sup>, Anna Gonzalez-Neira<sup>20</sup>, Anna H. Wu<sup>5</sup>, Annika Lindblom<sup>21</sup>, Anthony Swerdlow<sup>22,23</sup>, AOCs Study Group & Australian Cancer Study (Ovarian Cancer)<sup>24</sup>, APCB BioResource<sup>25,26</sup>, Argyrios Ziogas<sup>27</sup>, Arif B. Ekici<sup>28</sup>, Barbara Burwinkel<sup>29,30</sup>, Beth Y. Karlan<sup>31</sup>, Børge G. Nordestgaard<sup>32</sup>, Carl Blomqvist<sup>33</sup>, Catherine Phelan<sup>16</sup>, Catriona McLean<sup>34</sup>, Celeste Leigh Pearce<sup>35</sup>, Celine Vachon<sup>36</sup>, Cezary Cybulski<sup>37</sup>, Chavdar Slavov<sup>38</sup>, Christa Stegmaier<sup>39</sup>, Christiane Maier<sup>40</sup>, Christine B. Ambrosone<sup>41</sup>, Claus K. Høgdall<sup>42</sup>, Craig C. Teerlink<sup>43</sup>, Daehee Kang<sup>44,45</sup>, Daniel C. Tessier<sup>46</sup>, Daniel J. Schaid<sup>47</sup>, Daniel O. Stram<sup>5</sup>, Daniel W. Cramer<sup>48</sup>, David E. Neal<sup>49,50</sup>, Diana Eccles<sup>51</sup>, Dieter Flesch-Janys<sup>52</sup>, Digna R. Velez Edwards<sup>53</sup>, Dominika Wokozorczyk<sup>37</sup>, Douglas A. Levine<sup>54</sup>, Drakoulis Yannoukakos<sup>55</sup>, Elinor J. Sawyer<sup>56</sup>, Elisa V. Bandera<sup>57</sup>, Elizabeth M. Poole<sup>58,59</sup>, Ellen L. Goode<sup>36</sup>, Elza Khusnutdinova<sup>60,61</sup>, Estrid Høgdall<sup>62,63</sup>, Fengju Song<sup>64</sup>, Fiona Bruinsma<sup>65</sup>, Florian Heitz<sup>66,67</sup>, Francesmary Modugno<sup>68,69,70</sup>, Freddie C. Hamdy<sup>71,72</sup>, Fredrik Wiklund<sup>73</sup>, Graham G. Giles<sup>65,74,75</sup>, Håkan Olsson<sup>76</sup>, Hans Wildiers<sup>77</sup>, Hans-Ulrich Ulmer<sup>78</sup>, Hardev Pandha<sup>10</sup>, Harvey A. Risch<sup>79</sup>, Hatef Darabi<sup>73</sup>, Helga B. Salvesen<sup>80,81</sup>, Heli Nevanlinna<sup>82</sup>, Henrik Gronberg<sup>73</sup>, Hermann Brenner<sup>11,12,83</sup>, Hiltrud Brauch<sup>12,84,85</sup>, Hoda Anton-Culver<sup>27</sup>, Honglin Song<sup>3</sup>, Hui-Yi Lim<sup>86</sup>, Iain McNeish<sup>87</sup>, Ian Campbell<sup>88,89</sup>, Ignace Vergote<sup>90</sup>, Jacek Gronwald<sup>37</sup>, Jan Lubinski<sup>37</sup>, Janet L. Stanford<sup>91,92</sup>, Javier Benítez<sup>20</sup>, Jennifer A. Doherty<sup>93</sup>, Jennifer B. Permuth<sup>16</sup>, Jenny Chang-Claude<sup>19</sup>, Jenny L. Donovan<sup>94</sup>, Joe Dennis<sup>1</sup>, Joellen M. Schildkraut<sup>95,96</sup>, Johanna Schleutker<sup>97,98</sup>, John L. Hopper<sup>99</sup>, Jolanta Kupryjanczyk<sup>9</sup>, Jong Y. Park<sup>16</sup>, Jonine Figueroa<sup>100</sup>, Judith A. Clements<sup>101</sup>, Julia A. Knight<sup>102,103</sup>, Julian Peto<sup>104</sup>, Julie M. Cunningham<sup>105</sup>, Julio Pow-Sang<sup>16</sup>, Jyotsna Batra<sup>101</sup>, Kamila Czene<sup>73</sup>, Karen H. Lu<sup>106</sup>, Kathleen Herkommer<sup>107</sup>, Kay-Tee Khaw<sup>108</sup>, kConFab Investigators<sup>109</sup>, Keitaro Matsuo<sup>110</sup>, Kenneth Muir<sup>111,112</sup>, Kenneth Offitt<sup>113,114</sup>, Kexin Chen<sup>64</sup>, Kirsten B. Moysich<sup>115</sup>, Kristiina Aittomäki<sup>116</sup>, Kunle Odunsi<sup>117</sup>, Lambertus A. Kiemeny<sup>118</sup>, Leon F.A.G. Massuger<sup>119</sup>, Liesel M. Fitzgerald<sup>65</sup>, Linda S. Cook<sup>120</sup>, Lisa Cannon-Albright<sup>121</sup>, Maartje J. Hoening<sup>122</sup>, Malcolm C. Pike<sup>123</sup>, Manjeet K. Bolla<sup>1</sup>, Manuel Luedeke<sup>40</sup>, Manuel R. Teixeira<sup>17,124</sup>, Marc T. Goodman<sup>125</sup>, Marjanka K. Schmidt<sup>126</sup>, Marjorie Riggan<sup>127</sup>, Markus Aly<sup>73,128</sup>, Mary Anne Rossing<sup>91,92</sup>, Matthias W. Beckmann<sup>129</sup>, Matthieu Moisse<sup>130</sup>, Maureen Sanderson<sup>131</sup>, Melissa C. Southey<sup>89</sup>, Michael Jones<sup>22</sup>, Michael Lush<sup>1</sup>, Michelle A. T. Hildebrandt<sup>132</sup>, Ming-Feng Hou<sup>133</sup>, Minouk J. Schoemaker<sup>22</sup>, Montserrat Garcia-Closas<sup>22,100</sup>, Natalia Bogdanova<sup>134</sup>, Nazneen Rahman<sup>135</sup>,

NBCS Investigators<sup>136</sup>, Nhu D. Le<sup>137</sup>, Nick Orr<sup>138</sup>, Nicolas Wentzensen<sup>100</sup>, Nora Pashayan<sup>3,139</sup>, Paolo Peterlongo<sup>140</sup>, Pascal Guénel<sup>141,142</sup>, Paul Brennan<sup>143</sup>, Paula Paulo<sup>17</sup>, Penelope M. Webb<sup>144</sup>, Per Broberg<sup>145</sup>, Peter A. Fasching<sup>129</sup>, Peter Devilee<sup>146</sup>, Qin Wang<sup>1</sup>, Qiuyin Cai<sup>147</sup>, Qiyuan Li<sup>148,149</sup>, Radka Kaneva<sup>150</sup>, Ralf Butzow<sup>82,151</sup>, Reidun Kristin Kopperud<sup>80,81</sup>, Rita K. Schmutzler<sup>152,153</sup>, Robert A. Stephenson<sup>154</sup>, Robert J. MacInnis<sup>65,74</sup>, Robert N. Hoover<sup>100</sup>, Robert Winqvist<sup>155</sup>, Roberta Ness<sup>156</sup>, Roger L. Milne<sup>65,74</sup>, Ruth C. Travis<sup>157</sup>, Sara Benlloch<sup>1</sup>, Sara H. Olson<sup>123</sup>, Shannon K. McDonnell<sup>47</sup>, Shelley S. Tworoger<sup>58,59</sup>, Sofia Maia<sup>17</sup>, Sonja Berndt<sup>100</sup>, Soo Chin Lee<sup>158,159</sup>, Soo-Hwang Teo<sup>160,161</sup>, Stephen N. Thibodeau<sup>47</sup>, Stig E. Bojesen<sup>32,162,163</sup>, Susan M. Gapstur<sup>164</sup>, Susanne Krüger Kjær<sup>62,165</sup>, Tanja Pejovic<sup>166</sup>, Teuvo L.J. Tammela<sup>167</sup>, the GENICA Network<sup>168</sup>, the PRACTICAL consortium<sup>169</sup>, Thilo Dörk<sup>170</sup>, Thomas Brüning<sup>171</sup>, Tiina Wahlfors<sup>97</sup>, Tim J. Key<sup>157</sup>, Todd L. Edwards<sup>172</sup>, Usha Menon<sup>13</sup>, Ute Hamann<sup>78</sup>, Vanio Mitev<sup>150</sup>, Veli-Matti Kosma<sup>173,174</sup>, Veronica Wendy Setiawan<sup>5</sup>, Vessela Kristensen<sup>175,176,177</sup>, Volker Arndt<sup>11</sup>, Walther Vogel<sup>178</sup>, Wei Zheng<sup>147</sup>, Weiva Sieh<sup>14</sup>, William J. Blot<sup>179</sup>, Wojciech Kluzniak<sup>37</sup>, Xiao-Ou Shu<sup>147</sup>, Yu-Tang Gao<sup>180</sup>, Fredrick Schumacher<sup>5</sup>, Matthew L. Freedman<sup>181,182</sup>, Andrew Berchuck<sup>127</sup>, Alison M. Dunning<sup>3</sup>, Jacques Simard<sup>183</sup>, Christopher A. Haiman<sup>5</sup>, Amanda Spurdle<sup>184</sup>, Thomas A. Sellers<sup>16</sup>, David J. Hunter<sup>6</sup>, Brian E. Henderson<sup>5</sup>, Peter Kraft<sup>6</sup>, Stephen J. Chanock<sup>100</sup>, Fergus J. Couch<sup>105</sup>, Per Hall<sup>73</sup>, Simon A. Gayther<sup>5</sup>, Douglas F. Easton<sup>1,3</sup>, Georgia Chenevix-Trench<sup>2</sup>, Rosalind Eeles<sup>4,185</sup>, Paul D.P. Pharoah<sup>1,3</sup>, and Diether Lambrechts<sup>130</sup>

## Affiliations

<sup>1</sup>Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK <sup>2</sup>Department of Genetics, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia <sup>3</sup>Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK <sup>4</sup>The Institute of Cancer Research, Sutton, UK <sup>5</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, CA, USA <sup>6</sup>Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, MA, USA <sup>7</sup>Australian Breast Cancer Tissue Bank (ABCTB) Investigators, Westmead Millennium Institute, University of Sydney, Sydney, Australia <sup>8</sup>Division of Urologic Surgery, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Boston, MA, USA <sup>9</sup>Department of Pathology and Laboratory Diagnostics, the Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland <sup>10</sup>The University of Surrey, Guildford, UK <sup>11</sup>Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany <sup>12</sup>German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany <sup>13</sup>Women's Cancer, Institute for Women's Health, University College London, London, UK <sup>14</sup>Department of Health Research and Policy - Epidemiology, Stanford University School of Medicine, Stanford, CA, USA <sup>15</sup>Karolinska Institutet, Department of Environmental Medicine, Division of Nutritional Epidemiology, Stockholm, Sweden <sup>16</sup>Department of Cancer

Epidemiology, Moffitt Cancer Center, Tampa, FL, USA <sup>17</sup>Department of Genetics, Portuguese Oncology Institute, Porto, Portugal <sup>18</sup>Sheffield Cancer Research, Department of Oncology, University of Sheffield, Sheffield, UK <sup>19</sup>Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany <sup>20</sup>Human Cancer Genetics Program, Spanish National Cancer Research Centre (CNIO) and Spanish National Genotyping Center (CEGEN), Madrid, Spain <sup>21</sup>Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden <sup>22</sup>Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK <sup>23</sup>Division of Breast Cancer Research, The Institute of Cancer Research, London, UK <sup>26</sup>Australian Prostate Cancer Research Centre, Institute of Health and Biomedical Innovation and School of Biomedical Science, Queensland University of Technology, Brisbane, Queensland, Australia <sup>27</sup>Department of Epidemiology, UCI Center for Cancer Genetics Research and Prevention, School of Medicine, University of California Irvine, Irvine, CA, USA <sup>28</sup>University Hospital Erlangen, Institute of Human Genetics, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany <sup>29</sup>Molecular Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany <sup>30</sup>Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany <sup>31</sup>Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA <sup>32</sup>Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark <sup>33</sup>Department of Oncology, Helsinki University Hospital and University of Helsinki, Helsinki, Finland <sup>34</sup>Anatomical Pathology, The Alfred Hospital, Melbourne, Victoria, Australia <sup>35</sup>Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA <sup>36</sup>Department of Health Science Research, Division of Epidemiology, Mayo Clinic, Rochester, MN, USA <sup>37</sup>International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland <sup>38</sup>Department of Urology, Alexandrovska University Hospital, Medical University, Sofia, Bulgaria <sup>39</sup>Saarland Cancer Registry, Saarbrücken, Germany <sup>40</sup>Department of Urology, University Hospital Ulm, Ulm, Germany <sup>41</sup>Roswell Park Cancer Institute, Buffalo, NY, USA <sup>42</sup>The Juliane Marie Centre, Department of Gynecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark <sup>43</sup>Division of Genetic Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA <sup>44</sup>Cancer Research Institute, Seoul National University, Seoul, Korea <sup>45</sup>Departments of Preventive Medicine and Surgery, Seoul National University College of Medicine, Seoul, Korea <sup>46</sup>McGill University and Génome Québec Innovation Centre, Montréal, Canada <sup>47</sup>Mayo Clinic, Rochester, MN, USA <sup>48</sup>Obstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital, Boston, MA, USA <sup>49</sup>Department of Oncology, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK <sup>50</sup>Cancer Research UK Cambridge Institute, Li Ka Shing Centre, Cambridge, UK <sup>51</sup>Faculty of Medicine, University of Southampton, Southampton, UK <sup>52</sup>University Medical Center Hamburg-Eppendorf, Institute of Occupational Medicine

and Maritime Medicine and Institute for Medical Biometrics and Epidemiology, Hamburg, Germany <sup>53</sup>Vanderbilt Epidemiology Center, Vanderbilt Genetics Institute, Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, TN, USA <sup>54</sup>Gynecology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA <sup>55</sup>Molecular Diagnostics Laboratory, NCSR Demokritos, Agia Paraskevi, Athens, Greece <sup>56</sup>Research Oncology, Guy's Hospital, King's College London, London, UK <sup>57</sup>Cancer Prevention and Control Program, Rutgers Cancer Institute of New Jersey, The State University of New Jersey, New Brunswick, NJ, USA <sup>58</sup>Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA <sup>59</sup>Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA <sup>60</sup>Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russia <sup>61</sup>Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa, Russia <sup>62</sup>Department of Virus, Lifestyle and Genes, Danish Cancer Society Research Center, Copenhagen, Denmark <sup>63</sup>Molecular Unit, Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark <sup>64</sup>Department of Epidemiology and Biostatistics, Key Laboratory of Cancer Prevention and Therapy, Tianjin, National Clinical Research Center of Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin, P.R. China <sup>65</sup>Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia <sup>66</sup>Department of Gynecology and Gynecologic Oncology, Kliniken Essen-Mitte/ Evang. Huysens-Stiftung/Knappschaft GmbH, Essen, Germany <sup>67</sup>Department of Gynecology and Gynecologic Oncology, Dr. Horst Schmidt Kliniken Wiesbaden, Wiesbaden, Germany <sup>68</sup>Department of Obstetrics, Gynecology and Reproductive Sciences, Division of Gynecologic Oncology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA <sup>69</sup>Department of Epidemiology, University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA, USA <sup>70</sup>Womens Cancer Research Program, Magee-Womens Research Institute and University of Pittsburgh Cancer Institute, Pittsburgh, PA, USA <sup>71</sup>Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK <sup>72</sup>Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford, UK <sup>73</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden <sup>74</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Victoria, Australia <sup>75</sup>Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Victoria, Australia <sup>76</sup>Departments of Cancer Epidemiology and Oncology, University Hospital, Lund, Sweden <sup>77</sup>Department of General Medical Oncology, University Hospitals Leuven, Leuven Cancer Institute, Leuven, Belgium <sup>78</sup>Frauenklinik der Stadtklinik Baden-Baden, Baden-Baden, Germany <sup>79</sup>Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA <sup>80</sup>Department of Gynecology and Obstetrics, Haukeland University Hospital, Bergen, Norway <sup>81</sup>Centre for Cancer Biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway <sup>82</sup>Department of Obstetrics and Gynecology, Helsinki University Hospital and University of Helsinki, Helsinki, Finland <sup>83</sup>Division of

Preventive Oncology, National Center for Tumor Diseases (NCT) and German Cancer Research Center (DKFZ), Heidelberg, Germany <sup>84</sup>Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany <sup>85</sup>University of Tübingen, Tübingen, Germany <sup>86</sup>Biostatistics Program, Moffitt Cancer Center, Tampa, FL, USA <sup>87</sup>Institute of Cancer Sciences, University of Glasgow, Wolfson Wohl Cancer Research Centre, Beatson Institute for Cancer Research, Glasgow, UK <sup>88</sup>Cancer Genetics Laboratory, Research Division, Peter MacCallum Cancer Centre, St Andrews Place, East Melbourne, Victoria, Australia <sup>89</sup>Department of Pathology, University of Melbourne, Parkville, Victoria, Australia <sup>90</sup>Department of Gynaecologic Oncology, Leuven Cancer Institute, University Hospitals Leuven, KU Leuven, Leuven, Belgium <sup>91</sup>Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA <sup>92</sup>Department of Epidemiology, University of Washington, Seattle, WA, USA <sup>93</sup>Department of Epidemiology, The Geisel School of Medicine at Dartmouth, Hanover, NH, USA <sup>94</sup>School of Social and Community Medicine, University of Bristol, Bristol, UK <sup>95</sup>Department of Community and Family Medicine, Duke University Medical Center, Durham, NC, USA <sup>96</sup>Cancer Control and Population Sciences, Duke Cancer Institute, Durham, NC, USA <sup>97</sup>Department of Medical Biochemistry and Genetics Institute of Biomedicine, University of Turku, Turku, Finland <sup>98</sup>BioMediTech, University of Tampere and FimLab Laboratories, Tampere, Finland <sup>99</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia <sup>100</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA <sup>101</sup>Australian Prostate Cancer Research Centre, Institute of Health and Biomedical Innovation and School of Biomedical Science, Queensland University of Technology, Brisbane, Queensland, Australia <sup>102</sup>Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada <sup>103</sup>Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada <sup>104</sup>Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK <sup>105</sup>Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA <sup>106</sup>Department of Gynecologic Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA <sup>107</sup>Department of Urology, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany <sup>108</sup>Clinical Gerontology Unit, University of Cambridge, Cambridge, UK <sup>109</sup>Department of Preventive Medicine, Kyushu University Faculty of Medical Science, Nagoya, Aichi, Japan <sup>111</sup>Institute of Population Health, University of Manchester, Manchester, UK <sup>112</sup>Warwick Medical School, University of Warwick, Coventry, UK <sup>113</sup>Clinical Genetics Research Lab, Department of Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA <sup>114</sup>Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA <sup>115</sup>Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY, USA <sup>116</sup>Department of Clinical Genetics, Helsinki University



Hospital and University of Helsinki, Helsinki, Finland <sup>117</sup>Department of Gynecological Oncology, Roswell Park Cancer Institute, Buffalo, NY, USA <sup>118</sup>Radboud University Medical Centre, Radboud Institute for Health Sciences, Nijmegen, Netherlands <sup>119</sup>Department of Gynaecology, Radboud University Medical Center, Nijmegen, Netherlands <sup>120</sup>Division of Epidemiology and Biostatistics, Department of Internal Medicine, University of New Mexico, Albuquerque, NM, USA <sup>121</sup>George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT, USA <sup>122</sup>Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, The Netherlands <sup>123</sup>Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA <sup>124</sup>Biomedical Sciences Institute (ICBAS), University of Porto, Porto, Portugal <sup>125</sup>Cancer Prevention and Control, Samuel Oschin Comprehensive Cancer Institute, and Community and Population Health Research Institute, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA <sup>126</sup>Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands <sup>127</sup>Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC, USA <sup>128</sup>Department of Clinical Sciences, Danderyds Hospital, Stockholm, Sweden <sup>129</sup>Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany <sup>130</sup>VIB Vesalius Research Center, KU Leuven, Leuven, Belgium <sup>131</sup>Department of Family and Community Medicine, Meharry Medical College, Nashville, TN, USA <sup>132</sup>Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA <sup>133</sup>Cancer Center and Department of Surgery, Chung-Ho Memorial Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan <sup>134</sup>Radiation Oncology Research Unit, Hannover Medical School, Hannover, Germany <sup>135</sup>Section of Cancer Genetics, The Institute of Cancer Research, London, UK <sup>137</sup>Cancer Control Research, British Columbia Cancer Agency, Vancouver, BC, Canada <sup>138</sup>Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, UK <sup>139</sup>Department of Applied Health Research, University College London, London, UK <sup>140</sup>IFOM, The FIRC Institute of Molecular Oncology, Milan, Italy <sup>141</sup>Environmental Epidemiology of Cancer, Center for Research in Epidemiology and Population Health, INSERM, Villejuif, France <sup>142</sup>University Paris-Sud, Villejuif, France <sup>143</sup>International Agency for Research on Cancer, Lyon, France <sup>144</sup>Population Health Department, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia <sup>145</sup>Department of Cancer Epidemiology, University Hospital, Lund, Sweden <sup>146</sup>Departments of Human Genetics and of Pathology, Leiden University Medical Center, Leiden, The Netherlands <sup>147</sup>Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, Nashville, TN, USA <sup>148</sup>Department of Medical Oncology, The Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA <sup>149</sup>Medical College of Xiamen University, Xiamen, China <sup>150</sup>Department of Medical Chemistry and Biochemistry, Molecular Medicine Center, Medical University, Sofia, Bulgaria <sup>151</sup>Department of Pathology, Helsinki University Hospital and University of Helsinki, Helsinki, Finland <sup>152</sup>Center



for Integrated Oncology (CIO) and Center for Hereditary Breast and Ovarian Cancer, University Hospital of Cologne, Cologne, Germany <sup>153</sup>Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany <sup>154</sup>Department of Surgery, University of Utah School of Medicine, Salt Lake City, UT, USA <sup>155</sup>Laboratory of Cancer Genetics and Tumor Biology, Cancer Research and Translational Medicine, Biocenter Oulu, University of Oulu, and Northern Finland Laboratory Centre, Oulu, Finland <sup>156</sup>The University of Texas School of Public Health, Houston, TX, USA <sup>157</sup>Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK <sup>158</sup>Department of Hematology-Oncology, National University Health System, Singapore, Singapore <sup>159</sup>Cancer Science Institute of Singapore, National University of Singapore, Singapore, Singapore <sup>160</sup>Cancer Research Initiatives Foundation, Sime Darby Medical Centre, Subang Jaya, Malaysia <sup>161</sup>University Malaya Cancer Research Institute, University Malaya Medical Centre, University Malaya, Kuala Lumpur, Malaysia <sup>162</sup>Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark <sup>163</sup>Copenhagen General Population Study, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark <sup>164</sup>Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA <sup>165</sup>Department of Gynaecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark <sup>166</sup>Department of Obstetrics and Gynecology, and Knight Cancer Institute, Oregon Health and Science University, Portland, OR, USA <sup>167</sup>Department of Urology, Tampere University Hospital and Medical School, University of Tampere, Tampere, Finland <sup>170</sup>Gynaecology Research Unit, Hannover Medical School, Hannover, Germany <sup>171</sup>Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany <sup>172</sup>Vanderbilt Epidemiology Center, Division of Epidemiology, Department of Medicine, Vanderbilt Genetics Institute, Vanderbilt University, Nashville, TN, USA <sup>173</sup>Department of Pathology and Forensic Medicine, Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland <sup>174</sup>Department of Pathology, Kuopio University Hospital, Kuopio, Finland <sup>175</sup>Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway <sup>176</sup>K.G. Jebsen Center for Breast Cancer Research, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway <sup>177</sup>Department of Clinical Molecular Biology, Oslo University Hospital, University of Oslo, Oslo, Norway <sup>178</sup>Institute of Human Genetics, University Hospital Ulm, Ulm, Germany <sup>179</sup>Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA <sup>180</sup>Shanghai Cancer Institute, Shanghai, China <sup>181</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA <sup>182</sup>The Eli and Edythe L. Broad Institute, Cambridge, MA, USA <sup>183</sup>Genomics Center, Centre Hospitalier Universitaire de Québec Research Center, Laval University, Québec City, Canada <sup>184</sup>Molecular Cancer Epidemiology Laboratory, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia <sup>185</sup>Royal Marsden National Health Service (NHS) Foundation Trust, London and Sutton, UK

## Acknowledgments

### AOCS Study Group & Australian Cancer Study (Ovarian Cancer)

The following authors are part of the AOCS Study Group & Australian Cancer Study (Ovarian Cancer): Georgia Chenevix-Trench and Penelope M. Webb. Other members of the AOCS (nonauthor contributors) are listed under Supplementary Acknowledgments.

### Australian Prostate Cancer BioResource (APCB) BioResource

The following author is affiliated with the APCB BioResource: Judith A. Clements.

### Kathleen Cunningham Foundation Consortium for research into Familial Breast cancer (kConFab)

The following authors are affiliated with kConFab: Jonathan Beesley, Georgia Chenevix-Trench, Graham G. Giles, and Roger L. Milne. Other members (non-author contributors) of kConFab are listed under Supplementary Acknowledgments.

### Norwegian Breast Cancer Study (NBCS) Investigators

The following author is affiliated with NBCS: Vessela Kristensen. Other members (non-author contributors) of NBCS are listed under Supplementary Acknowledgments.

### The GENICA Network

The following authors are affiliated with the GENICA network: Hiltrud Brauch, Thomas Brüning, and Ute Hamann. Other members (nonauthor contributors) of the GENICA network are listed under Supplementary Acknowledgments.

### The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium

The following authors are affiliated with the PRACTICAL consortium: Rosalind A. Eeles, Douglas Easton, ZSofia Kote-Jarai, Ali Amin Al Olama, Sara Benlloch, Kenneth Muir, Graham G. Giles, Melissa C. Southey, Liesel M. Fitzgerald, Henrik Gronberg, Fredrik Wiklund, Markus Aly, Brian E. Henderson, Fredrick Schumacher, Christopher A. Haiman, Johanna Schleutker, Tiina Wahlfors, Teuvo LJ Tammela, Børge G. Nordestgaard, Tim J. Key, Ruth C. Travis, David E. Neal, Jenny L. Donovan, Freddie C. Hamdy, Paul Pharoah, Nora Pashayan, Kay-Tee Khaw, Janet L. Stanford, Stephen N. Thibodeau, Shannon K. McDonnell, Daniel J. Schaid, Christiane Maier, Walther Vogel, Manuel Luedeke, Kathleen Herkommer, Adam S. Kibel, Cezary Cybulski, Dominika Wokozorczyk, Wojciech Kluzniak, Lisa Cannon-Albright, Hermann Brenner, Aida k. Dieffenbach, Volker Arndt, Jong Y. Park, Thomas Sellers, Hui-Yi Lim, Chavdar Slavov, Radka Kaneva, Vanio Mitev, Jyotsna Batra, Amanda Spurdle, Manuel R. Teixeira, Paula Paulo, Sofia Maia, Hardev Pandha, Agnieszka Michael, and Andrzej Kierzek. Other members (nonauthor contributors) of the PRACTICAL consortium are listed under Supplementary Acknowledgments.

### **Financial Support**

The Breast Cancer Association Consortium (BCAC), the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL), and the Ovarian Cancer Association Consortium (OCAC) that contributed breast, prostate, and ovarian cancer data analyzed in this study were in part funded by Cancer Research UK [C1287/A10118 and C1287/A12014 for BCAC; C5047/A7357, C1287/A10118, C5047/A3354, C5047/A10692, and C16913/A6135 for PRACTICAL; and C490/A6187, C490/A10119, C490/A10124, C536/A13086, and C536/A6689 for OCAC]. Funding for the Collaborative Oncological Gene-environment Study (COGS) infrastructure came from: the European Community's Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-223175), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, and C8197/A16565), the US National Institutes of Health (CA128978) and the Post-Cancer GWAS Genetic Associations and Mechanisms in Oncology (GAME-ON) initiative (1U19 CA148537, 1U19 CA148065, and 1U19 CA148112), the US Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund [with donations by the family and friends of Kathryn Sladek Smith (PPD/RPCI. 07)]. Additional financial support for contributing studies is documented under Supplementary Financial Support.

## References

1. Henderson BE, Feigelson HS. Hormonal carcinogenesis. *Carcinogenesis*. 2000; 21:427–33. [PubMed: 10688862]
2. U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2012 Incidence and Mortality Web-based Report. [Internet]. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2015. [cited 2015 Sep 22]. Available from: [www.cdc.gov/uscs](http://www.cdc.gov/uscs)
3. Sellers TA, Potter JD, Rich SS, Drinkard CR, Bostick RM, Kushi LH, et al. Familial clustering of breast and prostate cancers and risk of postmenopausal breast cancer. *J Natl Cancer Inst*. 1994; 86:1860–5. [PubMed: 7990161]
4. Tung K-H, Goodman MT, Wu AH, McDuffie K, Wilkens LR, Nomura AMY, et al. Aggregation of ovarian cancer with breast, ovarian, colorectal, and prostate cancer in first-degree relatives. *Am J Epidemiol*. 2004; 159:750–8. [PubMed: 15051584]
5. Cerhan JR, Parker AS, Putnam SD, Chiu BC, Lynch CF, Cohen MB, et al. Family history and prostate cancer risk in a population-based cohort of Iowa men. *Cancer Epidemiol Biomarkers Prev*. 1999; 8:53–60. [PubMed: 9950240]
6. Amundadottir LT, Thorvaldsson S, Gudbjartsson DF, Sulem P, Kristjansson K, Arnason S, et al. Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. *PLoS Med*. 2004; 1:e65. [PubMed: 15630470]
7. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet*. 2015; 47:373–80. [PubMed: 25751625]
8. Kuchenbaecker KB, Ramus SJ, Tyrer J, Lee A, Shen HC, Beesley J, et al. Identification of six new susceptibility loci for invasive epithelial ovarian cancer. *Nat Genet*. 2015; 47:164–71. [PubMed: 25581431]
9. Al Olama AA, Kote-Jarai Z, Berndt SI, Conti DV, Schumacher F, Han Y, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nat Genet*. 2014; 46:1103–9. [PubMed: 25217961]
10. Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat Genet*. 2013; 45:371–84. 384e1–2. [PubMed: 23535731]
11. Stevens KN, Fredericksen Z, Vachon CM, Wang X, Margolin S, Lindblom A, et al. 19p13.1 is a triple-negative-specific breast cancer susceptibility locus. *Cancer Res*. 2012; 72:1795–803. [PubMed: 22331459]
12. Bolton KL, Tyrer J, Song H, Ramus SJ, Notaridou M, Jones C, et al. Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat Genet*. 2010; 42:880–4. [PubMed: 20852633]
13. Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, et al. A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. *Nat Genet*. 2010; 42:885–92. [PubMed: 20852631]
14. Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet*. 2013; 45:392–8. 398e1–2. [PubMed: 23535733]
15. Kote-Jarai Z, Olama AAA, Giles GG, Severi G, Schleutker J, Weischer M, et al. Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. *Nat Genet*. 2011; 43:785–91. [PubMed: 21743467]
16. Sakoda LC, Jorgenson E, Witte JS. Turning of COGS moves forward findings for hormonally mediated cancers. *Nat Genet*. 2013; 45:345–8. [PubMed: 23535722]
17. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet*. 2013; 14:483–95. [PubMed: 23752797]
18. Couch FJ, Nathanson KL, Offit K. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. *Science*. 2014; 343:1466–70. [PubMed: 24675953]

19. Nichols KE, Malkin D, Garber JE, Fraumeni JF, Li FP. Germ-line p53 mutations predispose to a wide spectrum of early-onset cancers. *Cancer Epidemiol Biomarkers Prev.* 2001; 10:83–7. [PubMed: 11219776]
20. Bhattacharjee S, Rajaraman P, Jacobs KB, Wheeler WA, Melin BS, Hartge P, et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am J Hum Genet.* 2012; 90:821–35. [PubMed: 22560090]
21. Han B, Sul JH, Eskin E, de Bakker PIW, Raychaudhuri S. A general framework for meta-analyzing dependent studies with overlapping subjects in association mapping. 2014 arXiv. [q-bio.QM]: 1304.8045v3.
22. Zernakova A, Stahl EA, Trynka G, Raychaudhuri S, Festen EA, Franke L, et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet.* 2011; 7:e1002004. [PubMed: 21383967]
23. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015; 348:648–60. [PubMed: 25954001]
24. Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet.* 2013; 45:1238–43. [PubMed: 24013639]
25. Corradin O, Scacheri PC. Enhancer variants: evaluating functions in common disease. *Genome Med.* 2014; 6:85. [PubMed: 25473424]
26. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. *Cell.* 2013; 155:934–47. [PubMed: 24119843]
27. He B, Chen C, Teng L, Tan K. Global view of enhancer-promoter interactome in human cells. *Proc Natl Acad Sci U S A.* 2014; 111:E2191–9. [PubMed: 24821768]
28. Corradin O, Saiakhova A, Akhtar-Zaidi B, Myeroff L, Willis J, Cowper-Salari R, et al. Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res.* 2014; 24:1–13. [PubMed: 24196873]
29. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* 2014; 159:1665–80. [PubMed: 25497547]
30. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell.* 2012; 148:84–98. [PubMed: 22265404]
31. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012; 40:D930–4. [PubMed: 22064851]
32. Gong J, Liu W, Zhang J, Miao X, Guo A-Y. lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucleic Acids Res.* 2015; 43:D181–6. [PubMed: 25332392]
33. Bhattacharya A, Ziebarth JD, Cui Y. PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. *Nucleic Acids Res.* 2014; 42:D86–91. [PubMed: 24163105]
34. Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinformatics.* 2012; 28:1797–9. [PubMed: 22513993]
35. Ashkenazi A, Dixit VM. Death receptors: signaling and modulation. *Science.* 1998; 281:1305–8. [PubMed: 9721089]
36. O'Connor L, Strasser A, O'Reilly LA, Hausmann G, Adams JM, Cory S, et al. Bim: a novel member of the Bcl-2 family that promotes apoptosis. *EMBO J.* 1998; 17:384–95. [PubMed: 9430630]
37. Bouillet P, Metcalf D, Huang DC, Tarlinton DM, Kay TW, Köntgen F, et al. Proapoptotic Bcl-2 relative Bim required for certain apoptotic responses, leukocyte homeostasis, and to preclude autoimmunity. *Science.* 1999; 286:1735–8. [PubMed: 10576740]
38. Lai AY, Wade PA. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat Rev Cancer.* 2011; 11:588–96. [PubMed: 21734722]

39. Zhai G, Teumer A, Stolk L, Perry JRB, Vandenput L, Coviello AD, et al. Eight common genetic variants associated with serum DHEAS levels suggest a key role in ageing mechanisms. *PLoS Genet.* 2011; 7:e1002025. [PubMed: 21533175]
40. Labrie F. Intracrinology. *Mol Cell Endocrinol.* 1991; 78:C113–8. [PubMed: 1838082]
41. Labrie F, Luu-The V, Labrie C, Bélanger A, Simard J, Lin S-X, et al. Endocrine and intracrine sources of androgens in women: inhibition of breast cancer and other roles of androgens and their precursor dehydroepiandrosterone. *Endocr Rev.* 2003; 24:152–82. [PubMed: 12700178]
42. Tworoger SS, Missmer SA, Eliassen AH, Spiegelman D, Folkert E, Dowsett M, et al. The association of plasma DHEA and DHEA sulfate with breast cancer risk in predominantly premenopausal women. *Cancer Epidemiol Biomarkers Prev.* 2006; 15:967–71. [PubMed: 16702378]
43. Key T, Appleby P, Barnes I, Reeves G. Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst.* 2002; 94:606–16. [PubMed: 11959894]
44. Arnold JT. DHEA metabolism in prostate: For better or worse? *Mol Cell Endocrinol.* 2009; 301:83–8. [PubMed: 19013497]
45. Takahashi H, Nakajima A, Sekihara H. Dehydroepiandrosterone (DHEA) and its sulfate (DHEAS) inhibit the apoptosis in human peripheral blood lymphocytes. *J Steroid Biochem Mol Biol.* 2004; 88:261–4. [PubMed: 15120419]
46. Charalampopoulos I, Tsatsanis C, Dermitzaki E, Alexaki V-I, Castanas E, Margioris AN, et al. Dehydroepiandrosterone and allopregnanolone protect sympathoadrenal medulla cells against apoptosis via antiapoptotic Bcl-2 proteins. *Proc Natl Acad Sci U S A.* 2004; 101:8209–14. [PubMed: 15148390]
47. Carmena M, Wheelock M, Funabiki H, Earnshaw WC. The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat Rev Mol Cell Biol.* 2012; 13:789–803. [PubMed: 23175282]
48. Samejima K, Platani M, Wolny M, Ogawa H, Vargiu G, Knight PJ, et al. The Inner Centromere Protein (INCENP) Coil Is a Single  $\alpha$ -Helix (SAH) Domain That Binds Directly to Microtubules and Is Important for Chromosome Passenger Complex (CPC) Localization and Function in Mitosis. *J Biol Chem.* 2015; 290:21460–72. [PubMed: 26175154]
49. Xu Z, Ogawa H, Vagnarelli P, Bergmann JH, Hudson DF, Ruchaud S, et al. INCENP-aurora B interactions modulate kinase activity and chromosome passenger complex localization. *J Cell Biol.* 2009; 187:637–53. [PubMed: 19951914]
50. Hilton JF, Shapiro GI. Aurora kinase inhibition as an anticancer strategy. *J Clin Oncol.* 2014; 32:57–9. [PubMed: 24043748]
51. Kabisch M, Lorenzo Bermejo J, Dünnebier T, Ying S, Michailidou K, Bolla MK, et al. Inherited variants in the inner centromere protein (INCENP) gene of the chromosomal passenger complex contribute to the susceptibility of ER-negative breast cancer. *Carcinogenesis.* 2015; 36:256–71. [PubMed: 25586992]
52. James LI, Barsyte-Lovejoy D, Zhong N, Krichevsky L, Korboukh VK, Herold JM, et al. Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nat Chem Biol.* 2013; 9:184–91. [PubMed: 23292653]
53. Cai Q, Zhang B, Sung H, Low S-K, Kweon S-S, Lu W, et al. Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. *Nat Genet.* 2014; 46:886–90. [PubMed: 25038754]
54. Metodiev MD, Spähr H, Loguerio Polosa P, Meharg C, Becker C, Altmueller J, et al. NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. *PLoS Genet.* 2014; 10:e1004110. [PubMed: 24516400]
55. Hirano T. Condensins: universal organizers of chromosomes with diverse functions. *Genes Dev.* 2012; 26:1659–78. [PubMed: 22855829]
56. Dávalos V, Suárez-López L, Castaño J, Messent A, Abasolo I, Fernandez Y, et al. Human SMC2 protein, a core subunit of human condensin complex, is a novel transcriptional target of the WNT signaling pathway and a new therapeutic target. *J Biol Chem.* 2012; 287:43472–81. [PubMed: 23095742]



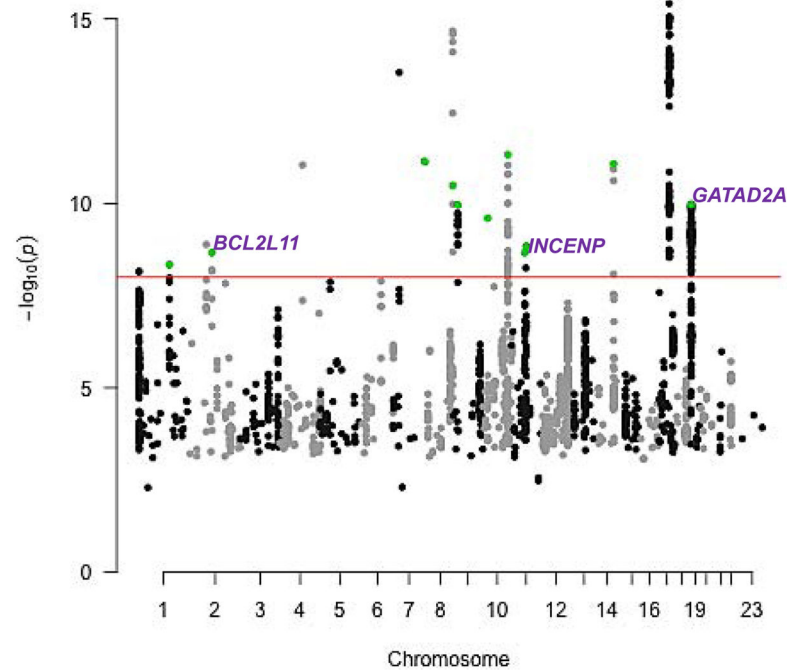
57. Murakami-Tonami Y, Kishida S, Takeuchi I, Katou Y, Maris JM, Ichikawa H, et al. Inactivation of SMC2 shows a synergistic lethal response in MYCN-amplified neuroblastoma cells. *Cell Cycle*. 2014; 13:1115–31. [PubMed: 24553121]
58. Wu N, Yu H. The Smc complexes in DNA damage response. *Cell Biosci*. 2012; 2:5. [PubMed: 22369641]
59. Wang D, Lu J, Tindall DJ. Androgens regulate TRAIL-induced cell death in prostate cancer cells via multiple mechanisms. *Cancer Lett*. 2013; 335:136–44. [PubMed: 23402817]
60. Keane MM, Ettenberg SA, Nau MM, Russell EK, Lipkowitz S. Chemotherapy augments TRAIL-induced apoptosis in breast cell lines. *Cancer Res*. 1999; 59:734–41. [PubMed: 9973225]
61. Lane D, Côté M, Grondin R, Couture M-C, Piché A. Acquired resistance to TRAIL-induced apoptosis in human ovarian cancer cells is conferred by increased turnover of mature caspase-3. *Mol Cancer Ther*. 2006; 5:509–21. [PubMed: 16546965]
62. Lewis-Wambi JS, Jordan VC. Estrogen regulation of apoptosis: how can one hormone stimulate and inhibit? *Breast Cancer Res*. 2009; 11:206. [PubMed: 19519952]
63. Zhang Y, Zhang B. TRAIL resistance of breast cancer cells is associated with constitutive endocytosis of death receptors 4 and 5. *Mol Cancer Res*. 2008; 6:1861–71. [PubMed: 19074831]
64. Allen JE, Kringsfeld G, Mayes PA, Patel L, Dicker DT, Patel AS, et al. Dual inactivation of Akt and ERK by TIC10 signals Foxo3a nuclear translocation, TRAIL gene induction, and potent antitumor effects. *Sci Transl Med*. 2013; 5:171ra17.
65. Leong S, Cohen RB, Gustafson DL, Langer CJ, Camidge DR, Padavic K, et al. Mapatumumab, an antibody targeting TRAIL-R1, in combination with paclitaxel and carboplatin in patients with advanced solid malignancies: results of a phase I and pharmacokinetic study. *J Clin Oncol*. 2009; 27:4413–21. [PubMed: 19652058]
66. Mendel G. Versuche über Pflanzenghybriden. *Verhandlungen Naturforschenden Vereines Brunn*. 1866; 4:3–47.
67. Freedman ML, Monteiro ANA, Gayther SA, Coetzee GA, Risch A, Plass C, et al. Principles for the post-GWAS functional characterization of cancer risk loci. *Nat Genet*. 2011; 43:513–8. [PubMed: 21614091]
68. Li C, Yang C, Gelernter J, Zhao H. Improving genetic risk prediction by leveraging pleiotropy. *Hum Genet*. 2014; 133:639–50. [PubMed: 24337655]
69. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 491:56–65. [PubMed: 23128226]
70. Bahcall OG. iCOGS collection provides a collaborative model. Foreword. *Nat Genet*. 2013; 45:343. [PubMed: 23535721]
71. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010; 26:2190–1. [PubMed: 20616382]
72. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*. 2015; 31:3555–7. [PubMed: 26139635]
73. Lin D-Y, Sullivan PF. Meta-analysis of genome-wide association studies with overlapping subjects. *Am J Hum Genet*. 2009; 85:862–72. [PubMed: 20004761]
74. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++ *PLoS Comput Biol*. 2010; 6:e1001025. [PubMed: 21152010]
75. Garber M, Guttman M, Clamp M, Zody MC, Friedman N, Xie X. Identifying novel constrained elements by exploiting biased substitution patterns. *Bioinformatics*. 2009; 25:i54–62. [PubMed: 19478016]
76. Matys V, Fricke E, Geffers R, Gössling E, Haubrock M, Hehl R, et al. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res*. 2003; 31:374–8. [PubMed: 12520026]
77. Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, Valen E, et al. JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res*. 2010; 38:D105–10. [PubMed: 19906716]



78. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489:57–74. [PubMed: 22955616]
79. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010; 26:2336–7. [PubMed: 20634204]
80. Nyholt DR. SECA: SNP effect concordance analysis using genome-wide association summary results. *Bioinformatics*. 2014; 30:2086–8. [PubMed: 24695403]
81. van Heyningen V, Bickmore W. Regulation from a distance: long-range control of gene expression in development and disease. *Philos Trans R Soc Lond B Biol Sci*. 2013; 368:20120372. [PubMed: 23650642]
82. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005; 102:15545–50. [PubMed: 16199517]
83. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet*. 2010; 11:843–54. [PubMed: 21085203]
84. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*. 2010; 38:W214–20. [PubMed: 20576703]

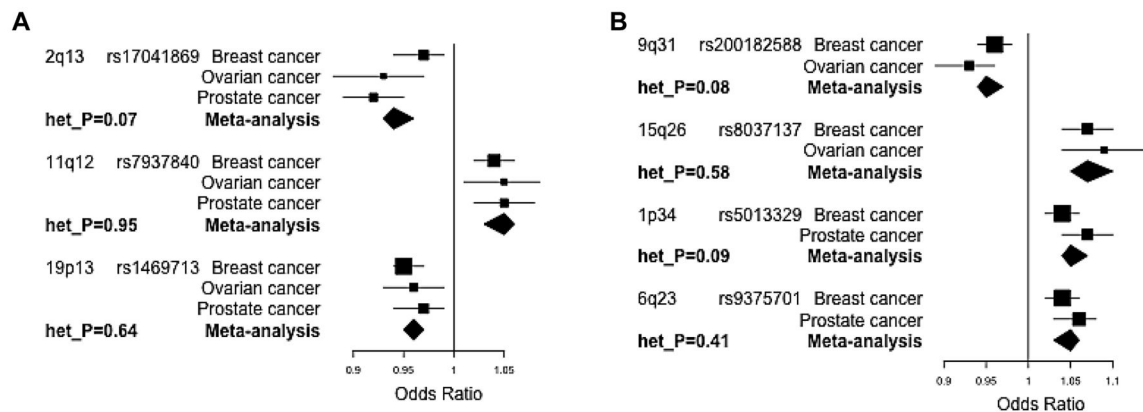
**Significance**

We demonstrate that combining large-scale genome-wide association meta-analysis findings across cancer types can identify completely new risk loci in common to breast, ovarian, and prostate cancer. We show that the identification of such cross-cancer risk loci has the potential to shed new light on the shared biology underlying these hormone-related cancers.



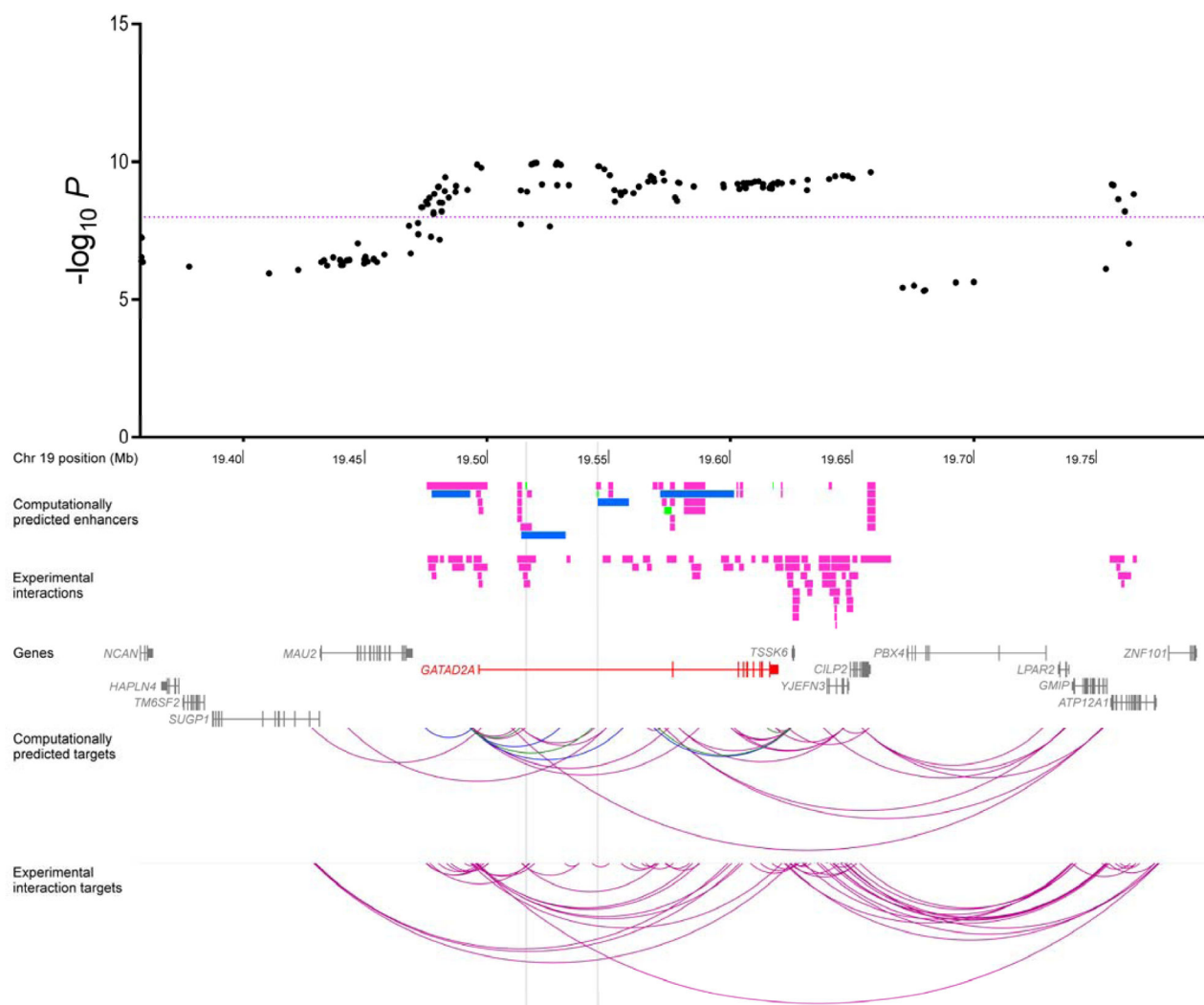
**Figure 1.**

Manhattan plot of results from the combined breast, ovarian, and prostate cancer meta-analysis. The black and gray dots represent the 2,231 variants nominally associated ( $P < 0.05$ ) with every cancer type individually that had the same direction of effect across all three cancers. The red line corresponds to a threshold of  $P = 10^{-8}$ . Eighteen independent loci were identified at this threshold. The green dots highlight index SNPs at 11 loci out of these 18 where model selection using ASSET confirmed contribution from all three cancer types to the association signal and that remained at  $P < 10^{-8}$  after adjusting for the controls shared between the breast and ovarian cancer studies. Gene names identify the three loci out of the 11 that were  $> 1$  Mb away from previously identified index SNPs for any of the three cancers.

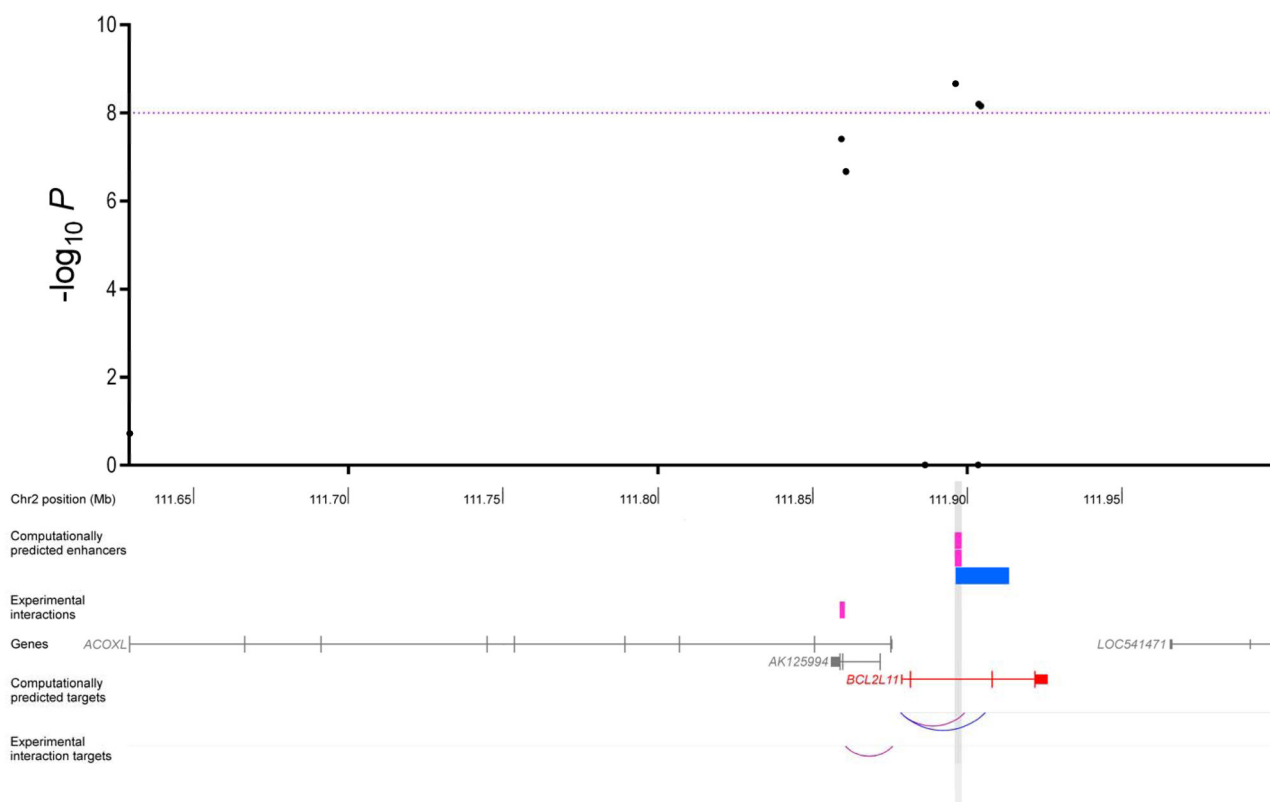


**Figure 2.**

Forest plots of odds ratio estimates for the new cross-cancer index SNPs (> 1 Mb from known index SNPs) for susceptibility to (A) breast, ovarian, and prostate cancer and (B) breast and ovarian cancer, and breast and prostate cancer. Error bars indicate 95% confidence intervals and het\_P is the *P*-value calculated from Cochran's *Q*-test for heterogeneity.



**Figure 3.** Regional association plot of results from the three-cancer meta-analysis for the rs1469713/19p13 breast, ovarian, and prostate cancer susceptibility locus. The black dots represent all variants nominally associated ( $P < 0.05$ ) with every cancer type individually that had the same direction of effect across all three cancers. The purple dashed line corresponds to a threshold of  $P = 10^{-8}$ . Tracks immediately below the regional association plot show the locations of enhancers in breast (pink), ovarian (green), and prostate (blue) cell types. Interactions derived from ChIA-PET experiments, which have only been assayed in breast cells, are labeled as experimental interactions. Where the same gene is predicted to be a target of enhancers that intersect with the same  $P < 10^{-8}$  SNP in all three cell types (or two for the 2q13 region), it is shown in red. All other genes in the region are in gray. The corresponding  $P < 10^{-8}$  SNP locations are marked by grey vertical stripes. The lower tracks show arcs between enhancers and target genes for both computationally predicted and experimentally derived interactions. Arc colors reflect the cell type in which the enhancer-promoter pair was identified.



**Figure 4.**

Regional association plot of results from the three-cancer meta-analysis for the rs17041869/2q13 breast, ovarian, and prostate cancer susceptibility locus. The black dots represent all variants nominally associated ( $P < 0.05$ ) with every cancer type individually that had the same direction of effect across all three cancers. The purple dashed line corresponds to a threshold of  $P = 10^{-8}$ . Tracks immediately below the regional association plot show the locations of enhancers in breast (pink), ovarian (green), and prostate (blue) cell types. Interactions derived from ChIA-PET experiments, which have only been assayed in breast cells, are labeled as experimental interactions. Where the same gene is predicted to be a target of enhancers that intersect with the same  $P < 10^{-8}$  SNP in all three cell types (or two for the 2q13 region), it is shown in red. All other genes in the region are in gray. The corresponding  $P < 10^{-8}$  SNP locations are marked by grey vertical stripes. The lower tracks show arcs between enhancers and target genes for both computationally predicted and experimentally derived interactions. Arc colors reflect the cell type in which the enhancer-promoter pair was identified.



New associations with a second cancer at known single-cancer risk loci<sup>a</sup>

Abbreviations: (E/R), (effect/reference) alleles; EAF: effect allele frequency.

 $f_{\text{Genotyped SNP}}$

Table 2

New cross-cancer loci identified at  $P < 10^{-8}$  that were over 1 Mb away from known index SNPs

Region, position <sup>a</sup>	Index SNP, (n), nearest gene	Alleles (E/R), EAF	Cancer Type	OR (95% CI)	P	ASSET model, $P_{het}$	$r^2c$
<b>A: From the three-cancer meta-analysis</b>							
<i>Associations with breast, ovarian and prostate cancer risk with the same direction of effect</i>							
2q13	rs17041869	A/G	Breast cancer	0.97 (0.94–0.99)	$7.1 \times 10^{-3}$	3-cancer	$1^d$
111896243	(3)	0.88	Ovarian cancer	0.93 (0.88–0.97)	$5.3 \times 10^{-4}$	0.07	$1^d$
	<i>BCL2L1</i>		Prostate cancer	0.92 (0.89–0.95)	$2.6 \times 10^{-6}$		$1^d$
			<b>Meta-analysis</b>	<b>0.94 (0.93–0.96)</b>	<b><math>5.1 \times 10^{-9}</math></b>		
11q12	rs7937840	T/C	Breast cancer	1.04 (1.02–1.06)	$3.6 \times 10^{-5}$	3-cancer	0.89
61893972	(1)	0.26	Ovarian cancer	1.05 (1.01–1.09)	$5.8 \times 10^{-3}$	0.95	0.90
	<i>INCENP</i>		Prostate cancer	1.05 (1.02–1.08)	$8.9 \times 10^{-4}$		0.89
			<b>Meta-analysis</b>	<b>1.05 (1.03–1.06)</b>	<b><math>5.0 \times 10^{-9}</math></b>		
19p13	rs1469713	A/G	Breast cancer	0.95 (0.94–0.97)	$9.9 \times 10^{-8}$	3-cancer	0.98
19528806	(89)	0.64	Ovarian cancer	0.96 (0.93–0.99)	$6.3 \times 10^{-3}$	0.64	0.98
	<i>GATAD2A</i>		Prostate cancer	0.97 (0.94–0.99)	$1.0 \times 10^{-2}$		0.95
			<b>Meta-analysis</b>	<b>0.96 (0.95–0.97)</b>	<b><math>3.4 \times 10^{-10}</math></b>		
<b>B: From the pairwise meta-analyses</b>							
<i>Associations with breast and ovarian cancer risk with the same direction of effect</i>							
9q31	rs200182588	G/GGC	Breast cancer	0.96 (0.94–0.98)	$1.9 \times 10^{-5}$	2-cancer	0.81
106856690	(15)	0.56	Ovarian cancer	0.93 (0.89–0.96)	$2.8 \times 10^{-6}$	0.08	0.82
	<i>SMC2</i>		<b>Meta-analysis</b>	<b>0.95 (0.94–0.97)</b>	<b><math>8.9 \times 10^{-9}</math></b>		
15q26	rs8037137	T/C	Breast cancer	1.07 (1.04–1.10)	$1.8 \times 10^{-7}$	2-cancer	0.98
91506637	(33)	0.86	Ovarian cancer	1.09 (1.04–1.14)	$2.1 \times 10^{-4}$	0.58	0.98
	<i>RCCD1</i>		<b>Meta-analysis</b>	<b>1.07 (1.05–1.10)</b>	<b><math>9.1 \times 10^{-10}</math></b>		
<i>Associations with breast and prostate cancer risk with the same direction of effect</i>							
1p34	rs5013329	T/C	Breast cancer	1.04 (1.02–1.06)	$7.8 \times 10^{-6}$	2-cancer	0.98
46815091	(218)	0.31	Prostate cancer	1.07 (1.04–1.10)	$1.4 \times 10^{-7}$	0.09	0.98
	<i>NSUN4</i>		<b>Meta-analysis</b>	<b>1.05 (1.04–1.07)</b>	<b><math>1.8 \times 10^{-11}</math></b>		

Region, position <sup>d</sup>	Index SNP, (n), nearest gene	Alleles (E/R), EAF	Cancer Type	OR (95% CI)	P	ASSET model, $P_{\text{het}}^b$	$r^2c$
6q23	rs9375701	T/C	Breast cancer	1.04 (1.02–1.06)	$3.6 \times 10^{-6}$	2-cancer	0.99
130384057	(53)	0.67	Prostate cancer	1.06 (1.03–1.08)	$1.5 \times 10^{-5}$	0.41	0.99
	<i>L3MBTL3</i>		Meta-analysis	<b>1.05 (1.03–1.06)</b>	<b><math>3.4 \times 10^{-10}</math></b>		

Abbreviations: n, Number of SNPs with  $P < 10^{-8}$  within 1 Mb of the index SNP; (E/R), (effect/reference) alleles; EAF: effect allele frequency.

<sup>a</sup>Build 37 coordinates.

<sup>b</sup>Cochran's  $Q$ -test for heterogeneity  $P$ -value.

<sup>c</sup>Imputation accuracy,  $r^2$ , in iCOGS European samples.

<sup>d</sup>Genotyped SNP.